LASER INDUCED TEMPERATURE JUMP INVESTIGATIONS OF FAST PROTEIN FOLDING DYNAMICS

By

LINLIN QIU

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2003

To Rongliang and Tianyue

ACKNOWLEDGMENTS

I would like to sincerely thank my advisor, Dr. Hagen, for three years of guidance and advice, for the days he spent teaching me experimental methodology and skills, the conferences and seminars he encouraged and supported me to attend, for the model he set for being a good experimentalist, and for all the fun he helped me find in doing protein folding research.

I am honored to have Dr. Cheng, Dr. Reitze, Dr. Rinzler, Dr. Roitberg, and Dr. Stanton as the members of my committee. I appreciate all the time they spent on my studies and all their valuable advice.

I would like to extend my appreciation to Dr. Zachariah and Alfred Chung for providing me the materials and to the nice people in machine shop and electronic shop for making me many high quality parts and devices.

I thank my colleagues, Suzette Pabit, OJ Ganesh, Priiyank Shukla and Geoff Gordon for their cooperation and help and thank my friends, Zhihong Chen, Xu Du and Minghan Chen, for sharing their experience and lab instruments and for many helpful discussions.

I wish to thank my husband Rongliang for all the support during the course of my research, all those helpful discussions, the knowledge and computer skills I learned from him, and all the time he spent checking the format of this dissertation.

I give my thanks to my parents, for their financial and moral support and for the time they spent taking care of my son for me so I can concentrate on my research.

Lastly, I wish to acknowledge my friends, Bob and Janis Jackson and Suzette Pabit for proofreading this dissertation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

LASER INDUCED TEMPERATURE JUMP INVESTIGATIONS OF FAST PROTEIN FOLDING DYNAMICS

By

Linlin Qiu

August 2003

Chair: Stephen J. Hagen
Major Department: Physics

Protein folding has a large parameter space, diverse mechanism, and multi-path kinetics. However, there are some common features many proteins share in their folding processes: all seem to fold at the rates much faster than the random conformation search, and all fold into the structures which have the highly regular motifs like α-helices, β-sheets and turns. Understanding how fast proteins can fold is one of the central issues in solving the protein folding problem.

Ultrafast folding kinetics had not been accessible until a few sub-millisecond probes were invented and applied lately. We constructed a laser induced temperature jump spectrometer which is a great utility in identifying the local structure and tertiary contact formation of proteins on the time scale from 10^{-8} to 10^{-3} s with time resolution of 10^{-9} s.

With this spectrometer we studied the fast folding mini-protein, TrpCage and a few short stable β-hairpins, the TrpZip series. Studying TrpCage was a major breakthrough it was a pioneer protein model which brought experiment and simulation very close: its structures measured by NMR and predicted by the molecular dynamics were amazingly alike. Our kinetic results showed that it folds

in 4 μs at room temperature which turned out to be the fastest ever known for protein-like molecules. Also this folding time constant is consistent with what was later on simulated by distributed computation. TrpZips are among the smallest and stablest polypeptide chains which form secondary structures. They are slightly different from each other based on structural stability and by forming various types of β-hairpins which are the minimum units of β tertiary structure. The β-hairpins form in the time range of 1–10 μs that confirms the theory that loop formation is controlled by the diffusion process ($\sim \mu$s).

We also investigated the kinetics of the protein chain collapse, a very controversial problem. By comparing the collapse of the foldable 104-residue protein cytochrome c and its unfoldable fragments F1-65 and F1-80, we concluded that the collapse of the protein molecule is not significantly different from those of unfoldable peptide chains. Burial of hydrophobic core and the presence of the interactions among chain residues and the interactions between amino acids and solvent molecules limit the collapse rate of a polypeptide chain on the time scale of the order of ten microseconds.

CHAPTER 1
GENERAL INTRODUCTION

1.1 Background

Sometimes physicists do care about the life sciences studying some problems like protein folding. Understanding protein folding helps people predict protein structures from amino acid sequences, design new drugs and understand why and how proteins misfold. Protein misfolding causes diseases like Madcow and Alzheimer's.

Twenty kinds of amino acids are the building blocks of protein molecules. They are connected by peptide bonds like a pearl necklace as shown in figure 1–1. An amino acid usually contains a carboxyl group, an amino group and a side chain—all bonded to the α-carbon atom. The peptide bond is an amide formed between the carboxyl group of one amino acid and the amino group of the next amino acid on the chain. Once the polypeptide chain is synthesized, it can fold spontaneously from a one-dimensional chain into an unique three-dimensional structure fitting into an extremely crowded cell, and becomes biologically active. This structure is defined as the native structure of a protein. The experimental studies by Anfinsen [1] in 1971 and his followers suggested that under the physiological conditions the native structure is solely determined by the amino acid sequence, no matter what folding pathway proteins experience.

This fascinating relationship between sequence and structure is a result of nature's evolution. This is also the key difference distinguishing protein molecules from homopolymers and synthesized heteropolymers. The spatial and temporal efficiencies of protein folding stimulate people from different disciplines to make efforts on revealing the inside story. In the last thirty years, a great development

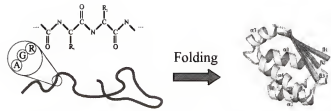


Figure 1–1: Protein folding scheme: A random polypeptide chain folds into a 3-D structure. A, R and G represent the amino acids: Alanine, Arginine and Glycine. The folded is the native structure of barstar — the 10kDa polypeptide inhibitor of the ribonuclease barnase.

has been achieved on establishing theories, carrying out computer simulations and collecting experimental data. However, some essential problems on folding kinetics still remain unclear and challenging:

1. How is a protein molecule able to fold spontaneously into its unique native structure without doing random search on all possible conformations from its unfolded state?
2. What are the preferred pathways by which proteins fold efficiently? Are intermediate states obligatory? What roles are they playing in protein folding?
3. How fast can a protein fold? What sets the folding time limit?
4. How different is the mechanism of protein folding from that of the packing of homopolymer chains, especially at the early folding stage?

Different from the ordinary chemical reactions in which usually a few high energy covalent bonds break and combine, protein folding generally involves a number of non-covalent interactions. The interactions include those among the amino acids of the protein molecules and those between protein molecules and the environment. This dramatically increases the complexity of this problem.

From a chemical point of view, there are mainly three folding driving forces: the hydrophobic effect, hydrogen bond, configurational entropy. Each kind of force involves the changing of energy. The difference in Gibbs free-energy between the

native state and the unfolded state of protein molecules is small ($\Delta G_{u \to f} \sim$10 to 70 kJ/mol) [2]. It is a delicate balance between two effects ($\Delta G = \Delta H - T\Delta S$): the enthalpy due to the non-covalent interactions like hydrophobic interaction and the configurational entropy due to the highly organized native structure. In addition, the external factors like temperature, pH and concentration of salt in solvent also influence protein folding processes. As a result, it is very difficult to predict the native structure for a protein molecule only based on its amino acid sequence information.

1.2 Hierarchical Protein Structure and Two-State Kinetics

In spite of the extremely large parameter space of folding, a number of small single-domain proteins and structure units are discovered to fold in a very simple mechanism: the two-state transition [3].

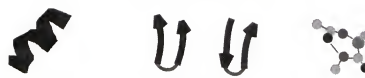
1.2.1 Protein Structure

Native protein structures are hierarchical as shown in figure 1–2: *primary structure* is the sequence of covalently linked amino acid residues; *secondary structure* is the regular way in which the polypeptide chain is arranged in space; *tertiary structure* is the biologically active, three-dimensional structure of an entire polypeptide; *quaternary structure* is an organization of two or more polypeptide chains into a multi-subunit protein.

The secondary structures like α-helices, β-sheets and turns are minimum units and highly regular motifs of protein structures. There is much evidence showing that the folding of large proteins begins with the formation of those secondary structures [4, 5]. A large body of work has been focused on the formation of secondary structures and the folding of the small single-domain proteins, polypeptide chains of a few helices, turns or loops [6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18].

Quaternary Structure

Tertiary Structure



Secondary Structure

-Ala-Glu-Val-Thr-Asp-Pro-Gly-

Primary Structure

Figure 1–2: Hierarchical structures of protein molecules. *Quaternary structure*: The four subunit arrangement of D-Xylose Isomerase is a complex, interwoven arrangement. *Tertiary structure*: (A) triosephosphate isomerase (B) hen egg lysozyme (C) myoglobin (D) chymotrypsin. *Secondary structure*: helix, β-sheet (parallel and anti-parallel) and turn. *Primary structure*: amino acid sequence

1.2.2 Two-State Transition

In a two-state folding system, there are only two kinds of stable molecular ensembles any time during the folding process: the unfolded(U) and the folded (F), as shown in equation (1.1). k_1 and k_{-1} are the folding and the unfolding rates.

$$U \overset{k_1 k_{-1}}{\rightleftharpoons} F \tag{1.1}$$

Many investigators in theoretical and simulation studies have been attracted to the two-state folding kinetics because of its simplicity. *Transition State Theory*(TST) is one of the simplest and most dominant models in the experimental literature on protein folding. It is a theory of rates of the elementary reactions which are assumed to occur in a single step and to pass through a single transition state. The transition state of more positive molar Gibbs energy between the reactants and the products would have an equal probability of forming the reactants and the products of the elementary reactions. The application of TST on protein folding is demonstrated in figure 1–3. The rate of the folding reaction is given by equation 1.2

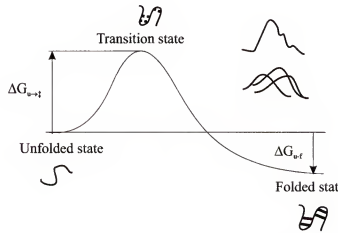


Figure 1–3: Transition state theory. $\Delta G_{u \to \ddagger}$ is the Gibbs energy difference between the unfolded state and the transition state. $\Delta G_{u \to f}$ is the energy difference between the the unfolded and folded. Two insets demonstrate the complexity and diversity of the transition state theory microscopically.

$$k_{u\to f} = (k_B T/h)\exp(-\Delta G_{u\to\ddagger}/(RT)), \tag{1.2}$$

where k_B is the Boltzmann constant, h is the Planck constant, T is temperature, R is the molar gas constant in kJ/mol, $\Delta G_{u\to\ddagger}$ is the Gibbs energy of activation.

Microscopically, systems may have more than one energy barrier along the folding pathway, but the highest one will be rate-limiting and its transition state is considered to be one for the overall reaction. On the other hand, there may be more than one pathway parallel to the direction of reaction coordinate. Each pathway has its own rate-limiting step but the apparent transition will have the averaged properties of all paths, weighted by the flux through each path [6].

For a two-state folder, given the environmental parameters such as temperature (T), pressure (P) or electric field strength (E), the populations of the folded and unfolded molecules keep a quasi-balance. Any perturbation on these parameters will trigger a redistribution of the population of molecules in each state. For a reversible two-state transition system, the perturbation will be followed by a relaxation of the concentrations of folded or unfolded molecules. This relaxation is treated as a single exponential function of time with observed rate constant $k_{\text{obs}} = k_1 + k_{-1}$. This is derived from equation 1.3 to equation 1.6 and based on equation 1.1 [19]:

$$\frac{\mathrm{d}[F]}{\mathrm{d}t} = k_1[U] - k_{-1}[F],\ \frac{\mathrm{d}[U]}{\mathrm{d}t} = k_{-1}[F] - k_1[U] \tag{1.3}$$

$$\frac{\mathrm{d}[F]}{\mathrm{d}t} = -(k_1 + k_{-1})[F] + k_1[UF] \tag{1.4}$$

$$[UF] \equiv [U] + [F] \tag{1.5}$$

$$[F](t) = C_1\exp(-k_{\text{obs}}t) + C_2. \tag{1.6}$$

$[F]$ and $[U]$ are the concentrations of molecules in folded and unfolded state, t is time, C_1 and C_2 are the constants which depends on the initial ($t = 0$) and final

($t >> \tau$) populations of folded molecules. The techniques capable of tracing this relaxation with a satisfactory time resolution can be used to probe protein folding kinetics.

Most two-state folding proteins and peptides are single-domain, of small size ~100 residues or less, and with simple structures of a few α helices, β strands or loops. However, they show a wide variation on the folding rates from microseconds to seconds. People believe that the folding of these proteins may represent the formation of the unit structures of large proteins or mimic the initial folding events which may set the tone for folding. Therefore these small proteins or structural units could be the ideal models both for theoretical and experimental studies. In addition, the molecular dynamics (MD) simulation on protein folding kinetics was just able to approach the microsecond time domain [20]. The proteins which fold in a few microseconds can provide the testing ground MD simulation is crying for [21, 22, 23].

1.3 Polymer Dynamics and Collapse

Recent studies show that there is more than one folding process even for two-state folders. Some fast relaxation events may precede or follow the rate-limiting barrier crossing phase [24, 25, 26, 27]. They are probably due to the fast chain collapse, protein molecule desolvation or equilibration processes on the time scale which is comparable to that of folding [28]. The mechanisms of these processes and the roles they play in protein folding have not been fully understood. These processes may reflect the general and rapid response of polypeptide chains to the change of environmental parameters such as solvent, temperature, etc. and may set the time limit for protein folding processes [29, 30].

Monte Carlo simulations on lattice and off-lattice models [31, 32, 33, 34] addressed whether the collapse-induced compactness of polypeptide chain enhances the formation of secondary structures. Maritan [33, 34] demonstrated that there is

a strong correlation between the fast folding, content of secondary structures and compactness of protein molecules. This suggests that the collapse could be a key factor to understand the spatial and temporal efficiency of protein folding. Here arise two questions:

1. How fast could a polypeptide chain collapse be?
2. Do protein molecules collapse in the same way as homopolymers?

In classical polymer dynamics [35], all segments of a polymer chain are considered to be Brownian particles whose motion in dilute solution is governed by the Smoluchowski or Langevin equation. The connection of consecutive segments, the interactions between solvent and polymer molecules and long range intramolecular interactions(excluded volume effect) are included in the potentials which generate the force field for the Brownian particles. In 1953 , Rouse pioneered in building such a model on an ideal chain. A few years later, the Zimm model [36] introduced the hydrodynamic interaction between the particles on the chain which led to the exciting agreement with experimental results on polymers. Zimm's model study on polymer dynamics may set the limit for the time scale of protein folding which is on the nanosecond or microsecond time domain depending on the size of molecules and the external conditions like the viscosity of solvent and temperature.

1.4 High Resolution Measurements on Structures and Kinetics

Sub-millisecond folding kinetics could not be identified until a number of high time resolution techniques were invented in the last ten years [6, 37] listed in table 1–1. Mixing, temperature jump, pressure jump and photolysis are ultrafast triggers based on concentration, temperature, pressure and photochemical reactions to shift folding equilibrium rapidly. The relaxation following ultrafast triggers can be monitored with fluorescence, infrared, resonance Raman and circular dichroism spectroscopies listed in table 1–2. Some equilibrium methods like nuclear magnetic

resonance (NMR) or electron paramagnetic resonance (EPR) can also provide kinetic information on a few hundred microseconds time scale [7, 38].

Table 1–1: Ultrafast triggers for protein folding kinetics.

Method	Probe time window	Structure resolution
Ultrafast Mixing	10 μs–∞	atomic
Laser Flash Photolysis	100 fs–1 ms	individual residue
Laser Induced Temperature Jump	1 ns–100 ms	individual residue
NMR Line broadening	100 μs–100 ms	atomic

Ultrafast mixing spectroscopy triggers and probes refolding of protein molecules directly by quickly diluting the denaturant or varying the conditions of solutions like pH or ligand and then monitoring the structure-related changes of the optical properties during folding process. There are two kinds of flow schemes in mixing experiments: stop-flow and continuous-flow. Stop-flow combined with circular dichroism and fluorescence spectroscopies is commercially available presently but its dead time is about a few milliseconds. The fastest mixing device is the continuous-flow system which can barely reach the microsecond time regime.

Table 1–2: Spectroscopic probes for protein folding kinetics.

Method	Sensitive to
Fluorescence	tertiary structure and environment
Infrared	strength of bonds, dipole moments, polarizability
Resonance Raman	vibration
Circular Dichroism	Secondary structures

Laser flash photolysis is the one which has the best time resolution ($\sim$ 100 fs) so far. It has been successfully used to study the kinetics of cytochrome c refolding, chain diffusion and some small conformation changes [39, 40]. In the cytochrome c refolding experiment, lasers were used to trigger the ultrafast protein refolding by photolyzing the carbon monoxide(CO) which is bonded to the heme of cytochrome c and the conformational changes following the protein-CO dissociation can be monitored by spectroscopies. Although the time resolution was limited by the

nanosecond CO rebinding event, photo-dissociation itself takes less than 100 femtoseconds. Its disadvantage is that this method only applies to a limited number of proteins in which the folding or unfolding can be triggered by photochemical reactions.

Most proteins change their structures in response to the changing of environmental parameters like temperature, pressure or dielectric coefficient. Based on this fact, laser or electric induced temperature jump, pressure jump and acoustic or dielectric relaxation of the nanosecond to microsecond time resolution are being used widely on protein folding kinetic studies.

1.5 Work We Have Done

We used Nd-YAG laser pulses at 1064 nm to pump stimulated Raman scattering of hydrogen gas (H_2). This generates near infrared pulses at 1890 nm. At this wavelength, aqueous solution absorbs light tremendously and releases energy as heat in tens of nanoseconds that sets the dead time for our temperature jump system. The probe system is designed for the proteins containing fluorescent residues which absorb the ultraviolet light at about 266 nm and emit the fluorescence with the center wavelength of 350 nm. The emission intensity of protein molecules can be sensitive not only to the environmental parameters like temperature and pH but also to proteins' tertiary structures according to the Förster transfer theory [41].

We carried out the ultrafast folding kinetic studies on a few two-state folders, TrpCage and the tryptophan zipper (TrpZips) series, with the temperature jump spectrometer we built. The small helical protein or those β-hairpins turned out to be the ideal models for the dynamic study of protein folding. Since the folding rates approach the diffusion time limit for the polypeptide chains, which is $\sim \mu$s, some interesting effects such as the breaking of equilibration assumption for the *two-state transition theory* or for the relaxation principle, begin to show up in the kinetic data of the fast folders like TrpZip2 under certain conditions. On the other

hand, because microsecond kinetics lately became accessible to molecular dynamics simulation, our results also provided the testing grounds for the simulation work done by Pande's group.

In addition, we addressed the ultrafast chain collapse problems by studying the collapse kinetics of wild type cytochrome c and its unfoldable fragments F1-65 and F1-80. We found out that the collapse kinetics of the unfoldable polypeptide chains is not significantly different from that of wild type protein and concluded that 10 μs collapse of the foldable and unfoldable polypeptide chains is probably controlled by the chain diffusion, desolvation process and intrachain interaction instead of by forming any specific native structure.

CHAPTER 2
LASER INDUCED TEMPERATURE JUMP SPECTROSCOPY

2.1 Introduction

Among all the fast techniques, nanosecond laser induced temperature jump spectroscopy attracts much interest because it is suitable for the kinetic studies on all kinds of fast folding proteins and peptides. That is based on the fact that proteins generally keep their stable native structures over a certain temperature range and unfold at the temperatures beyond it. We can trigger either folding or unfolding of protein molecules by raising temperature quickly.

It was first recognized by Eigen and his colleagues in the 1960s [42, 43] that the rapid kinetic studies can be carried out by perturbing an equilibrium system by the temperature jump and observing the system relaxing from one equilibrium state to a new one after the perturbation. This basic scheme is shown in figure 2–1, using the Förster energy transfer fluorescence spectroscopy as an example. The fluorescent intensity of a fluorophore (F) on a polypeptide chain is mainly affected by two factors: temperature and the physical distance from its fluorescence light emission quencher (Q) somewhere on the same peptide chain. Suppose that the system is at an initial equilibrium state at the temperature T_{initial}, and its fluorescence intensity is at a certain level. At $t = 0$, the temperature of the solution jumps up to T_{final} by ΔT. The fluorescence signal responds instantly ($\sim$ ns) by dropping to a lower level because of the intrinsic negative temperature dependence of the fluorophores, but it takes a much longer time for the protein molecule to change conformation and for the fluorophore to move away from its quencher. This process accompanies an increasing of fluorescence intensity according to the Förster

energy transfer theory which will be discussed in more details in section 2.1.3. The single exponential relaxation indicates a two-state transition.

Several temperature jump experiments have been developed such as P.A. Thompson's at the National Institute of Health, B. Nölting's at Cambridge University, M. Gruebele's at the University of Illinois at Urbana-Champaign and R.B. Dyer's at Los Alamos National Laboratory. These systems share the same idea but are different from each other by the experimental layout and details such as the heating light sources, probe methods.

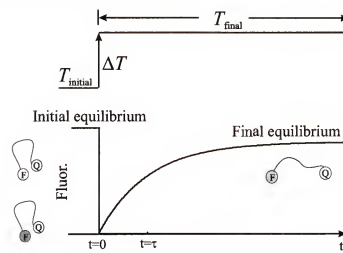


Figure 2–1: The demonstration of the temperature jump technique using the Förster energy transfer fluorescence spectroscopy as a probe. The top figure is the temperature jump scheme. The bottom one is the probed fluorescence signal demonstrating the folding kinetics of protein molecules.

The first fast protein collapse was observed by Gruebele in 1996: Apomyoglobin initially collapses into a compact state within 20 μs [44]. In 1999, 60 μs collapse of a 100-residue protein cytochrome c at room temperature was fully resolved by Hagen and Eaton in NIH [45]. The temperature jump with infrared probe spectroscopy turned out to be of much advantage in studying the secondary structure formation and the helix-coil transition. It was found out that α-helix forms in a few hundred nanoseconds [8, 46]. In addition, the sub-millisecond refolding of a cold-unfolded protein *Bovine β-lactoglobulin A* was discovered by

Nölting in 1996 [47]. These pioneer works unveiled the mystery of the early folding processes of proteins and drew great attention to this much unexplored territory.

Here we used 1064 nm infrared pulses to trigger the stimulated Raman scattering (SRS) of hydrogen gas (H_2). The first Stokes' line of H_2 is at 1890 nm ($\Delta\nu = 4108\ \mathrm{cm}^{-1}$) at which aqueous solutions have a tremendous absorption and release the energy as heat to induce a temperature jump of solutions in a few nanoseconds. The protein conformation dynamics is probed by the fluorescence Förster energy transfer spectroscopy.

2.1.1 Features of *T*-Jump Spectroscopy

Advantages. There are quite a few advantages for using the near infrared (IR) laser pulses to induce the temperature jump. Firstly, IR pulses can directly heat up the aqueous solutions in a few nanoseconds, there is no need for the extra reagents such as the dyes in some laser-induced temperature jump spectrometers and salt in electrical temperature jump spectrometers. The potential problems caused by these reagents are eliminated. Secondly, only a moderate amount of IR energy is necessary to get a reasonable temperature jump because the observing volume is determined by the diameter of the probe beam that could be confined into a small size ~1 mm or less. In addition, the pulse widths of heating and probe laser pulses are more compatible with each other and also with the response time of protein molecules [48].

Disadvantages. The usage of the temperature jump spectrometer is limited by the following factors:

1. The upper limit of the observable time window of the temperature jump apparatus cannot be as infinitely large as the fast mixers because the transiently raised temperature (T_{final} in figure 2–1) decays due to the thermal diffusion of solutions after a period of time. The T_{final}-stable span is from

t ~ns to about hundreds of μs or to tenths of seconds which relies on the thermal properties of sample solutions and the geometry of the sample holder.

2. Rapid temperature jump in a small region introduces the thermal lensing, photo-acoustic effect and cavitation which interfere with probing pulses and distort probed signals. Those effects are minimized at a temperature around 4°C and become obviously worse at the temperatures higher than 60°C [49].

2.1.2 Fluorescence Spectroscopy and Förster Energy Transfer

Many protein molecules contain fluorescent amino acids e.g. Tryptophan (Trp), Tyrosine (Tyr) and Phenylalanine (Phe) naturally or by mutation. The fluorescence characteristics of these amino acids are shown in table 2–1 where λ_{max} is the wavelength at which the substance has the maximum absorption or maximum fluorescence emission, ε_{max} is the extinction coefficient in $\mathrm{mol}^{-1}\mathrm{cm}^{-1}$ and ϕ_F is the fluorescence quantum yield which is equal to the ratio of photons emitted to photons absorbed by the system [41].

Table 2–1: Fluorescence characteristics of the three amino acids.

Substance	Conditions	Absorption		Fluorescence	
		λ_{max} (nm)	$\varepsilon_{max} \times 10^{-3}$	λ_{max} (nm)	ϕ_F
Tryptophan	H_2O, pH 7.0	280	5.6	348	0.20
Tyrosine	H_2O, pH 7.0	274	1.4	303	0.14
Phenylalanine	H_2O, pH 7.0	257	0.2	282	0.04

The fluorescence quantum yield, ϕ_F, is the result of the competition between the intrinsic fluorescence rate constant (k_F) with the processes including internal conversion (k_{ic}), intersystem crossing (k_{is}) and quenching [$k_q(Q)$].

$$\phi_F = k_F/[k_F + k_{ic} + k_{is} + k_q(Q)] \tag{2.1}$$

In internal conversion, the excitation energy of fluorophores is lost by collision with solvent molecules or by dissipation through internal vibration modes. The rate of the internal conversion increases as the temperature increases. In the

intersystem crossing process, the excited singlet state is converted into an excited triplet state through the normally forbidden spin exchange. The triplet state will then convert to the ground state either by phosphorescence or internal conversion. Among the various quenching types, the Förster fluorescence energy transfer is a probabilistic process based on the dipole-dipole interaction between suitable donor and acceptor molecules. The energy transfer efficiency (E) is the function of the distance between the donor and acceptor which is derived as follows.

The efficiency of energy transfer is defined in equation 2.2 as the function of transition rate k_T and the life time of the donor in the absence of the acceptor τ_D.

$$E = k_T/(k_T + 1/\tau_D) \tag{2.2}$$

$$k_T = (1/\tau_D)(R_0/R)^6 \tag{2.3}$$

By substituting k_T in equation 2.2 with equation 2.3, we rewrite E as follow,

$$E = \frac{R_0^6}{R_0^6 + R^6}, \tag{2.4}$$

R_0 is the characteristic separation distance (Förster distance) where the probability of the energy transfer is 50%.

$$R_0 = 9.7 \times 10^3 (J\kappa^2 n^{-4} \phi_D)^{1/6} \text{cm} \tag{2.5}$$

where

$$J = \int \varepsilon_A(\nu) f_D(\nu) \nu^{-4} \, d\nu \tag{2.6}$$

J is the overlap integral of the donor emission and acceptor absorption spectra; f_D is the normalized fluorescence emission of the donor; n is the refractive index of the solvent; ϕ_D is the fluorescence quantum yield of donor in the absence of the acceptor and κ^2 is a geometrical factor that depends on the relative orientation of donor and acceptor dipoles.

The Förster energy transfer from donor to acceptor results in the decreasing of the donor's fluorescence emission. The transfer efficiency is solely determined by the distance between the donor and acceptor as in equation 2.4. Figure 2–2 demonstrates the kinetic studies on protein cytochrome c by the Förster energy transfer spectroscopy. Figure 2–2 (A) [1] is the cartoon which shows the dipole-dipole interaction of the fluorescence donor, tryptophan and the acceptor, Heme at His18. Figure 2–2 (B) shows the kinetic results by the ultrafast mixing experiment [24]: when the protein molecule folds and becomes compact, the tryptophan fluorescence decreases due to the increase of the Förster energy transfer efficiency.

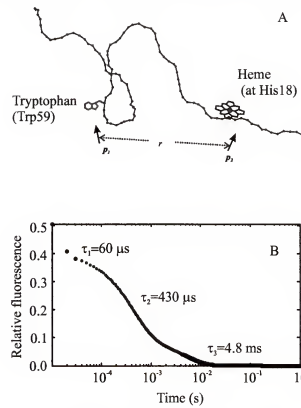


Figure 2–2: Application of the Förster energy transfer theory on cytochrome c. (A) Diagram of the cytochrome c chain with the fluorophore–Trp at No.59 and its quencher–Heme at His18. (B) Refolding kinetics of cytochrome c by the ultrafast mixing experiment.

[1] Provided by S.J. Hagen

2.2 System Buildup and Characterization

2.2.1 Goal and System Composition

The T-jump apparatus is composed of the following parts as shown in figure 2–3:

1. infrared heating pulses generation.
2. ultraviolet probe pulses generation.
3. sample holder.
4. temperature control and monitor system.
5. signal detection.
6. data collection and analysis.

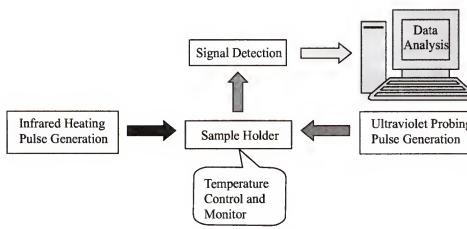


Figure 2–3: Diagram of the T-jump system.

An optimized system should have a widely accessible observing time window/long T_{final}-stable time span, very low background noise and robust temperature control and detection.

2.2.2 Infrared Heating Pulse Generation

We use Continuum Surelite I-10 Q-switched Nd:YAG laser to produce pulses (at λ=1064 nm, 2 Hz, with 5–7 ns pulse width) to pump a Raman cell filled with hydrogen gas. The first Stokes line of H_2 Raman scattering is at 1890 nm ($\Delta\nu$ = 4108 cm^{-1}) at which the absorption coefficient of water is about 30 cm^{-1}.

The Raman shifter output as the function of Q-switch[2] time delay of 1064 nm laser pulses is plotted in figure 2–4.

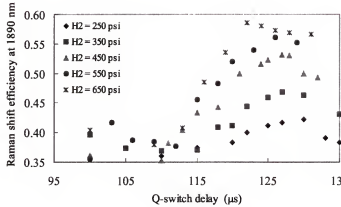


Figure 2–4: Raman shift efficiency at 1890 nm as the function of H_2 pressure and the Q-switch delay time of the infrared laser.

Studies suggested that adding some helium gas into the Raman cell increases the conversion efficiency. However according to figure 2–5 we show here that the mixture of He and H_2 does not produce a higher portion of output energy compared to pure H_2.

After the Raman cell, 1890 nm pulses reach the sample holder from two opposite directions as shown in figure 2–6, the layout of T-jump optics. This two-heating-pulse scheme helps improve the uniformity of T_{final} distribution inside the sample holder. The total energy of 10 mJ/pulse at 1890 nm generates the

[2] Q-switch is a device used to quickly change the Q value (Q is defined as $2\pi\times$ (average energy stored in the resonator)/(energy dissipated per cycle)) of an optical resonator . It is used in the optical resonator of a laser to prevent lasing action until a high level of inversion (optical gain and energy storage) is achieved in the lasing medium. When the switch rapidly increases the Q of the cavity, a short pulse of high intensity light is generated.

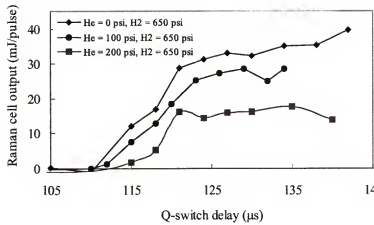


Figure 2–5: Raman shifter output as the function of the IR Q-switch time delay with the various pressures of He. The frequency of the 1064 nm input pulse is 10 Hz and the pressure of H_2 is 650 psi.

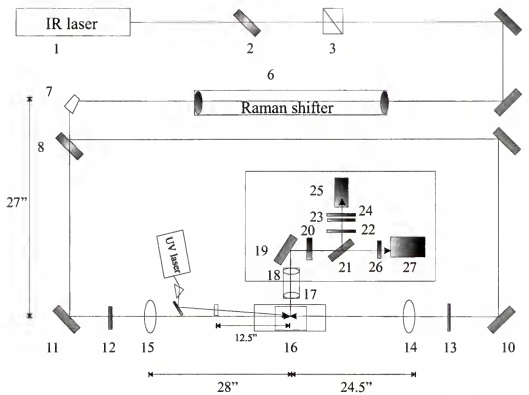
temperature jump ~20°C for aqueous solutions with a focus spot size 1 mm × 0.5 mm.

2.2.3 Ultraviolet Probe System

Ultraviolet probe pulses are generated by Continuum Minilite I-1 Nd:YAG Q-switched laser (at λ=266 nm, 2 Hz, with 5 ns width). UV beam is focused on the same spot as infrared beams are as shown in figure 2–6. The detailed optical layout in demonstrated in figure 2–7. The UV pulse intensity is monitored by a photodiode and recorded during the experiments for the signal intensity drift correction (see Software) and improving signal-to-noise ratio.

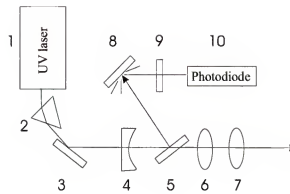
2.2.4 Sample Holder Design and Temperature Control

The temperature control and measurement are the critical issues in sample holder design. Experiment requires that the equilibrium temperature T_{initial} of the sample should be fully controlled and the temperature jump amplitude ΔT and T_{final} accurately measured. The system needs to work well within a wide temperature range.



1. Continuum Surelite Q-switched Nd:YAG laser
2. TFP polarizer
3. quarter-wave plate
4. mirror newport 10-D20 DM8
5. mirror newport 10-D20 DM8
6. raman shifter 1m long with 600 psi H2 with two lenses f=50cm AR coated NIR
7. Pellin broca prism
8. CVI beam splitter 2010nm R60%
9. CVI mirror IR 2010nm TM-1037-45UNP
10. CVI mirror IR 2010nm TM-1037-45UNP
11. CVI mirror IR 2010nm TM-1037-45UNP
12. Schott glass filter type 59880
13. Schott glass filter type 59880
14. lens f=50cm KPX118 BK7 DCK
15. Lens f=50cm KPX118 BK7 DCK
16. Temperature sample holder
17. MItutoyo objective f=200 mm
18. CVI mirror 266 nm Y4-1025-0
19. mirror
20. iris
21. flip mirror
22. neutral density filters
23. WG320 filter 1mm thick
24. Schott glass filter type Ug5
25. Hamamatsu R1166 photomultiplier
26. 3mm UG5 filter
27. CCD camera

Figure 2-6: T-jump setup diagram.



1. Continuum Minilite UV laser
2. prism
3. UV mirror
4. silica lens f= −50mm
5. silica window
6. silica lens f=100mm
7. silica lens f=50cm
8. reflecting plate with teflon tape
9. Schott WG320
10. photo-diode

Figure 2–7: UV optical layout.

Equilibrium temperature control and measurement.. Figure 2–8 shows the diagram of the sample holder. The main frame is made of aluminum because of its good thermal conductivity ($\lambda_{Al}(27°C) = 2.37$ W/cm-K). A rectangular silica tube ($0.1 \times 1 \times 30$ mm I.D.) is held by an aluminum jacket which keeps the temperature of the silica tube the same as the holder base. Protein solution flows through the rectangular tube. There are three windows on the jacket allowing IR and UV beams to be focused on the sample inside the tube and the fluorescence signals to come out. The thermocouple on the jacket, labeled with a dot in the figure, measures the temperature of the sample holder.

The temperature of the sample holder is controlled by a thermoelectric chip (TEC) beneath with an electronic feedback control circuit panel MPT (Wavelength Electronic MPT-5000). The equilibrium temperature can be stabilized in $T \pm 0.1°C$ within the range between 0°C and 100°C.

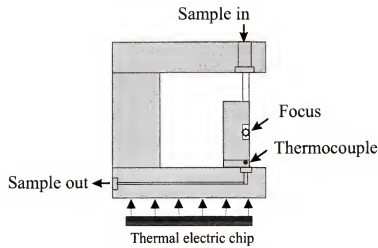


Figure 2–8: Sample holder diagram.

Theoretical limit for T_{final} duration. Without the thermal diffusion of sample solution, the laser induced temperature jump is given by equation (2.7) [48].

$$\delta T(r,z,t) = k/\rho c_v \int_0^1 I(r,z,t)\,\mathrm{d}t' \tag{2.7}$$

where $I(r,z,t)$ is the intensity profile of the heating laser pulse at time t, and $k/\rho c_v$ is determined by the properties of sample solution: ρ is the density of solution, c_v is the heat capacity, and k is the absorption cross section per unit length at the wavelength of the heating laser pulses.

In fact, the duration of the raised temperature, T_{final}, is influenced by the thermal diffusion process which is governed by the diffusion equation. In this case, the diffusion equation is constrained by the boundary conditions which are depending on the geometry of the sample holder. If the diameter of the heating spot is much smaller than the length of the sample holder and the sample holder can be treated as an infinite cylinder, we should expect a Gaussian function for the temperature profile. The temperature at the origin follows:

$$\delta T(t) = \delta T_{\max}\left[1 - \exp(-r^2/4\kappa t)\right] \tag{2.8}$$

where r is the radius of the cylinder and κ is the thermal diffusivity in cm^2/s. For the aqueous solution inside 100 μm×100 μm×30 mm ultraviolet transparent silica fiber, it roughly takes about 5 ms for the temperature at origin to decay by 5% with $\kappa = 1.43 \times 10^{-3}$ cm^2/s at 293K. This sets the theoretical upper limit for the observing time window of our T-jump system.

Reference sample for ΔT measurement. N-Acetyl-Trp-Amide (NATA) is the reference sample we use to optimize the alignment of the UV probe beam and IR heating beams and to calculate the amplitude of the temperature jump. The intrinsic temperature dependence of the fluorescence of Trp is shown in figure 2–9 measured by a fluorometer (Jasco FP-750). Its fluorescence intensity changes as a function of temperature:

$$F = \begin{cases} F_0 \exp(-0.022T) & \text{if } 50°\text{C} \geq T \geq 0°\text{C} \\ F_0 \exp(-0.025T) & \text{if } T > 50°\text{C} \end{cases} \tag{2.9}$$

Optical alignment: ΔT maximization. We first pump NATA solution into the sample holder and fire the UV pulse 1 μs later than the IR pulse. Then adjust the optics to maximize the fluorescence signal difference with and without IR pulses. The fluorescence signal is monitored by the Labview program, *alignmentsave-data.vi* [3] .

ΔT and T_{final} duration determination. Once the optics is optimized, we precede every protein data collection with an experiment on the NATA reference sample under the same condition by using the Labview program *tjumpRMS.vi* [4] . The typical reference data is shown in figure 2–10.

[3] see Appendix B–1 and B–2 for the front panel and the block diagram.

[4] see Appendix B–3 to B–8 for details.

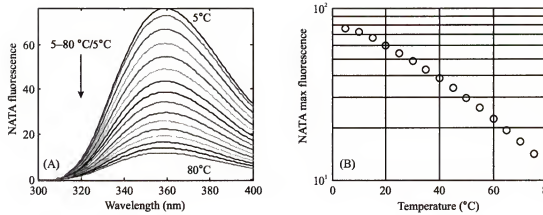


Figure 2–9: Temperature dependence of the fluorescence of NATA in 100 mM phosphate buffer pH 7.0. (A) Wavelength scan curves at different temperatures. (B) Maximum values of the each curve in (A) as the function of temperature.

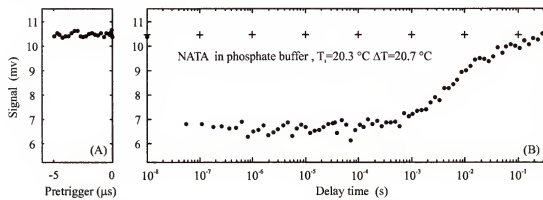


Figure 2–10: Temperature jump calibration data with NATA at 20 μM in 100 mM phosphate buffer pH 7.0. (A) and (B) Fluorescence signals of NATA before and after the temperature jump at $t = 0$.

Figure 2–10 includes two parts: pretrigger data ($t \leq 0$) and post trigger data ($t > 0$). Before the temperature jump, Trp fluorescence signal is about 10 mV at initial temperature of 20°C. When $t = 0$, the fluorescence intensity drops about 40% due to the 20-degree temperature jump. The following unchanged signal indicates that the T_{final} remains same within a millisecond until the heated volume of solution begins to diffuse out and Trp fluorescence relaxes to the pretrigger level accordingly with a time constant $\tau \sim$500 ms.

From this kinetic data on NATA, We obtain the fluorescence signal level $F(t = 0-)$, $F(t = 0+)$ and can calculate the temperature jump amplitude ΔT according to equation 2.9.

Furthermore, it takes ~30 ns for NATA molecules to respond to the infrared heating pulses that determines the dead time of T-jump system. As a result, this system has a more than four decade time window (from 30 ns to 1 ms) and a nanosecond time resolution for protein folding kinetics studies.

2.2.5 Signal Detection, Amplification and Measurement

The fluorescence signals are detected by a photomultiplier (PMT) (HAMAMATSU R1166 D-type ϕ 23.5×130 mm) and amplified 10 times by the amplifier (HAMAMATSU C6438) before they are measured by the digital oscilloscope (Tektronix, TDX3032). To keep the PMT working in a linear region, we use some neutral density filters in front of PMT to control the signal level within 100 mV – 200 mV at the voltage of PMT power supply of about 320 V. The characterization of the PMT detection linearity is shown in figure 2–11.

The data collection and the signal averaging are performed by an oscilloscope. We chose RMS as the measurement mode in order to make the signal less sensitive to the oscillating electronics noises coming from the Q-switches of the lasers. The output of the oscilloscope is then sent to a computer for the data analysis.

The noise in the T-jump experiment mainly comes from the following sources: the shot to shot stability of the IR and UV laser pulses, the sample flow stability and a few kinds of electronic noises.

The intensity and fluctuation of UV laser pulses as the function of the Q-switch time delay is shown in figure 2–12. The UV intensity increases with the Q-switch time delay when $t_{Q\_switch} < 150$ μs and decreases when $t_{Q\_switch} \geq 150$ μs. Meanwhile the STD/mean is at a stable level when $t_{\text{Q\_switch}} \geq 110$ μs, which is about 10%. Accordingly the UV induced Trp fluorescence will fluctuate unless

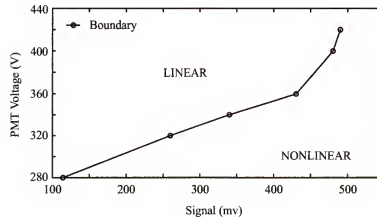


Figure 2–11: Photomultiplier linearity. The linear and nonlinear detection regions are separated by the boundary line with circles.

the excitation energy is high enough to saturate the absorption as in figure 2–13. Increasing the excitation energy is one way to suppress the fluctuation of fluorescence signals. On the other hand, high UV intensity will cause the instant decay of the fluorescence signal during a measurement. This problem could be solved by increasing the pumping velocity of the sample fluid. Figure 2–14 shows that the NATA fluorescence becomes stable at flowrate higher than 0.1 ml/hr with the UV beam focus size $\sim$ 50 μm. We chose to raise the UV pulse energy to a certain value which barely reaches the saturation threshold, $\sim$ 100 mV, measured by the photo-diode while keeping the flow rate as low as 0.1–0.5 ml/hr to avoid using a large amount of sample each time.

The fluctuation of ΔT due to the pulse-to-pulse infrared energy fluctuation is another source of noise. Because the stability of IR laser output becomes worse when t_{Q_switch} >125 μs, all of our T-jump experiments have been carried out with t_{Q_switch} below that value. Besides we found that this kind of fluctuation is randomly distributed temporally therefore it can be reduced by the signal averaging as shown in figure 2–15.

The laser beam pointing fluctuation has been overcome by moving the light sources as close as possible to the sample holder and changing the sample holder

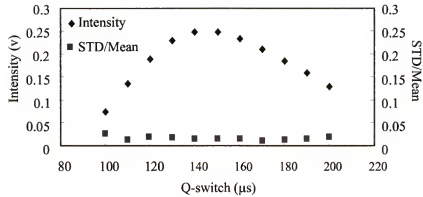


Figure 2–12: UV intensity versus UV laser Q-switch time delay. The diamonds represent the UV pulse intensities and the squares are standard deviation versus the mean UV pulse intensity.

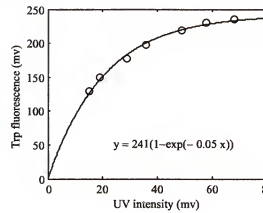


Figure 2–13: Trp fluorescence intensity versus the UV excitation energy while the sample flowrate is at 0.1 ml/hr and the diameter of the UV focus spot is ∼50 μm. The solid line is the single exponential function fit.

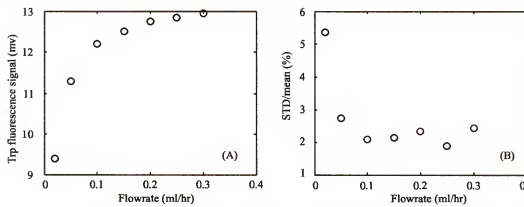


Figure 2–14: Trp fluorescence intensity (A) and fluctuation (B) versus the flowrate. The diameter of the UV focus spot is ∼50 μm.

from a square silica capillary into a rectangular silica tube with the wider sides facing the IR and UV beams.

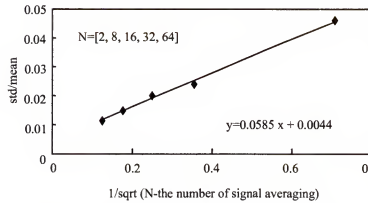


Figure 2–15: Intensity fluctuation of the infrared pulses (standard deviation/intensity) measured with an InGaAs photodiode at IRQT 125 μs. N is the number of the signal averaging. The solid line is the linear fit.

In addition, the electronic noise due to the other instruments was diminished by electrical shielding and signal averaging. Consequently the noise-to-signal ratio has been controlled within 2%.

2.2.6 Software

There are two kinds of software being used in the T-jump system: the Labview programs for computer to set parameters for instruments and to acquire data from them (*alignmentsavedata.vi* and *tjumpRMS.vi* by L. Qiu) and the Matlab programs for calculating temperature jump (*tjumpcal.m* by L. Qiu) and for data analysis (*Tjump_proc_1.m* by S.J.Hagen).

Labview programs for instruments. Controlling the time difference between IR and UV laser pulses is the key role the Labview programs play in the T-jump system. This involves four instruments: one function generator (Tektronix CFG280), one pulse generator (DG535, Stanford research system Inc.) and the power supplies of the IR and UV lasers. The pulse-timing control scheme is shown

in figure 2–16. The function generator generates 10 Hz TTL pulses to trigger the infrared laser flashlamp and the pulse generator DG535. The IR laser won't fire until its Q-switch is triggered by a pulse which comes from channel A of DG535. The output pulses from channel B, C of DG535 are respectively used to trigger the UV laser flashlamp and the UV laser Q-switch. The time difference between the IR flashlamp and Q-switch and the time difference between the UV flashlamp and Q-switch are set to be 110–125 μs and 110 μs. Channel D is set to be 415 ms later than the trigger for the DG535, so all channels A, B, C, D can send out pulses every 500 ms or at frequency 2 Hz. While the time delay between the IR and UV output pulses, t, varies during the experiments. The system measures the fluorescence signal at each t once or more times depending how many times signal needs to be averaged. A fluorescence signal value, an UV pulse intensity signal and the corresponding time, t, constitute one data unit for a data file.

tjumpRMS.vi is the primary Labview program to control the pulse timing and data collection. It has a front panel shown in appendix B–3 which allows users to specify the instruments for the computer to communicate with, set parameters for these instruments and provides the data collecting options such as the time range, the number of data points, the distribution of data points in linear or in logarithmic time scales. It plots data in real time and saves data automatically after each experiment in two ASCII files: one is a three-column matrix of time, fluorescence intensity and UV intensity and the other includes all critical parameters in an experiment and the description of the sample used in the experiment. Its working flowchart is in figure 2–17 and the block diagram is listed in Appendix B–4 to B–8.

In addition, there are two more Labview programs *alignmentsavedata.vi* and *DG535.vi* used to optimize the optical alignment, check sample flow stability and

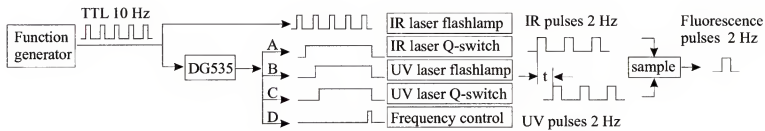


Figure 2–16: Pulse timing control scheme.

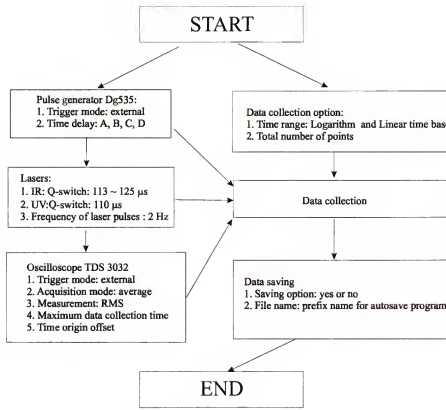


Figure 2–17: Flowchart of *tjumpRMS.vi*.

noise level etc. before experiments. Their front panels and block diagrams are listed in Appendix 6–7.

Matlab programs for data analysis. *Tjump_proc_1* is the main Matlab program for the data analysis. Its working sequence is as following,

1. Loads data saved by *tjumpRMS.vi*.
2. Corrects intensity drift using two options: UV intensity drift correction ($y = at + b$, a, b are fitting parameters, y and t are UV intensity signal and the corresponding time) only or along with sample intensity drift correction ($F = ct_{\text{pretrigger}} + d$, c, d are fitting parameters, F and $t_{\text{pretrigger}}$ are fluorescence intensity signal when $t < 0$ and the corresponding time).
3. Removes the "way off" data points due to the bubbles in solution or the cavitation effect during an experiment.

4. Fits data into the function

$$F = \begin{cases} F_0 & \text{if } t < 0; \\ a + b\exp(-t/\tau) & \text{otherwise,} \end{cases}$$

where F_0 is the averaged pretrigger signal level, a and b are fitting parameters and τ the relaxation time constant.

5. Calculates F_0, a, b, τ, and uncertainties of these parameters.

6. Plots data in linear time scale.

tjumpcal.m first loads NATA kinetic data and corrects for the drift like *tjump_proc_1.m*. It averages the pretrigger fluorescence signal and gets $F(t = 0-)$, then fits the posttrigger signal ($0 < t <$ 1ms) with the function $F = F_0 + b\,t$ which represents a slight cooling process of the sample within the observing time window due to the sometime imperfect optical alignment and obtains the first posttrigger level $F(t = 0+)$. It finally uses $F(t = 0-)$ and $F(t = 0+)$ to calculate the temperature jump amplitude based on the equation 2.9.

2.3 Summary and Future Plan

In our lab, a temperature jump spectrometer is built up with the main features as follow:

- Applicable to the proteins which: absorb ~266 nm, emit 320–375 nm
- Accessible observing time window: 30 ns–1 ms
- Time resolution: 1 ns
- Working temperature range: 0–100°C
- Equilibrium temperature stability: ±0.1°C
- Temperature jump calibration: according to the intrinsic fluorescence of Trp
- Temperature jump accuracy: ±1–2°C
- Noise to signal ratio: 1–2%
- Signal detection linearity: 100%
- Sample consumption rate: 0.1–0.5 ml/hr

- Instrument control program: *tjumpRMS.vi*
- Data analysis program: *tjump_proc_1.m*

In the future, we will research on an alternative way to calibrate the temperature jump by detecting the temperature jump induced change of the optical properties of samples such as the index of the refraction or light absorption at a certain wavelength so that we can measure the size of the temperature jump instantaneously. Also a further improvement on the signal-to-noise ratio should be achieved by correcting data according to the fluctuation of the infrared pulse energy.

CHAPTER 3
HELICES, HAIRPINS AND FAST FOLDING PROTEINS

3.1 Introduction

Cooperativity is one of the most fundamental concepts in protein folding studies. It states that some small-single-domain proteins fold like two-state systems with only the folded and unfolded states are stable while all partly folded or intermediate states are not. Since 1991, when chymotrpsin inhibitor 2 (CI2), a small helical protein was first proved experimentally to fold by two-state kinetics [50], many studies have been focused on this [51, 52, 53, 54, 55, 56, 57, 58].

Most of two state folders are single-domain proteins with a few secondary structure units like α-helices, β-sheets, and turns as shown in figure 1–2 but they show a wide variation in folding rates from s^{-1} to μs^{-1} which are mainly dependent on the topology of their native states [55]. With the development of ultrafast probes, more and more interesting kinetic events have been observed: α-helices form in a few hundred nanoseconds while β-hairpins ten times slower, $\sim \mu s$ [8, 9, 11]; some two-state folding proteins fold accompanied by some extra events like fast chain contraction although they may not produce structured intermediate states [26, 59]; some proteins fold much faster than theoretical models predicted like TrpCage [60]; when the folding is so fast as to approach its physical time limit, even the minimum structure formation deviates from the two-state kinetics as the TrpZip hairpin series we study here.

We carried out the kinetic studies on an ultrafast helical mini-protein, TrpCage and a few β-hairpins, TrpZips using our T-jump spectrometer [60]. The experimental results indicate that TrpCage is a typical two-state folding protein-like molecule that folds in 4 μs at room temperature, which became a record fast folding speed to

date. Three TrpZips fold in 1–10 μs which is consistent with the previous studies on β-hairpins [9]. It is not totally clear what governs the folding rates of these fast folders and what contributes to their activation energy barriers. The microsecond kinetics of TrpCage and TrpZips provides the test ground not only for theoretical models but also for the molecular dynamics simulation. We found some simulation studies [21] are in good agreement with our measurements.

3.2 TrpCage

3.2.1 Background

TrpCage is a helical mini-protein synthesized by Anderson's group [61]. It has twenty residues: NLYIQ WLKDG GPSSG RPPPS [1] . Structurally it contains one α-helix from residues 2 to 8, a 3_{10}-helix from 11 to 14 and a compact hydrophobic core where three proline residues (Pro-12, Pro-18, Pro-19) and a glycine (Gly-11) packed against the aromatic side chains of Tyr-3 and Trp-6 as shown in figure 3–1 [2] . It is known as one of the smallest protein-like molecules.

It has attracted a lot of interest from the simulation community because of its small size and good structure stability. Roitberg and coworkers predicted its structure from an all atom, fully unrestrained folding simulation. The predicted structure turned out to be amazingly close to the NMR resolved one [62]. Pande's group in Stanford University predicted the folding kinetics by distributed computation and showed that TrpCage may fold in a time constant between 1.5 μs and 6.9 μs at 27°C [21] which will prove to be consistent with the experimental results in the following sections.

---

[1] Each letter represents an amino acid. Also see Appendix A–1 for details

[2] Provided by A. Roitberg

3.2.2 Methods and Results

Equilibrium. We use the circular dichroism (CD) spectrometer to obtain the secondary structure content of TrpCage. Circular dichroism is the difference in absorbance between left and right circularly polarized light: $\Delta A = A_L - A_R$. An asymmetric molecule like a helical protein molecule exhibits circular dichroism which has two minima in the far-UV spectrum with one of them at 222 nm as shown in figure 3–2. CD is not only used to quantify the secondary structure, it is also capable of examining the change of the secondary structure content due to the environmental factors such as temperature or binding events.

On the other hand, the fluorescence of protein molecules is often used as an indicator of the compactness of the molecules. TrpCage is fluorescent due to the tryptophan (Trp, W) residue in its amino acid sequence at the position No.6. The intensity of TrpCage fluorescence decreases when the temperature increases, but this decreasing trend is not as much as a free tryptophan as shown in figure 3–3. This suggests that the thermally unfolding process of the TrpCage molecule enhances the tryptophan fluorescence emission that balances the negative intrinsic temperature dependent of Trp fluorescence ($\sim -2.2\%/°\text{C}$).

By the CD and fluorescence measurements, we obtained the two unfolding fraction curves which perfectly overlap each other as shown in figure 3–4. This indicates that the equilibrium data is independent of the probes satisfying one of the criteria for the two-state transition model [50].

Kinetics. The kinetic data of the TrpCage were obtained by using the T-jump spectroscopy. The sample was dissolved in 0.1 M potassium phosphate buffer pH7.0 at the concentration $\sim$50 μM. As shown in figure 3–5, at $t = 0$, the temperature jump from 11°C to 22°C causes a drop of the TrpCage fluorescence signal because of the T-dependence of Trp fluorescence. After the T-jump, the populations of the folded and unfolded TrpCage molecules will relax to the new

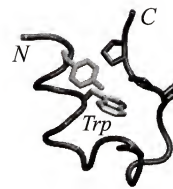


Figure 3–1: TrpCage structure.

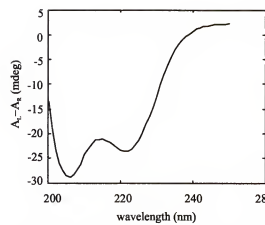


Figure 3–2: CD spectrum of TrpCage.

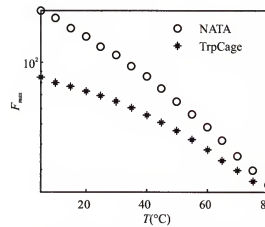


Figure 3–3: Temperature dependence of TrpCage and NATA fluorescence.

equilibrium levels at the temperature, $T_{\text{final}} = 22°\text{C}$. This population relaxation of the two species fits to the single exponential function very well as the solid line shown in figure 3–5.

Using the relationship between τ_{obs}, the folding rate, k_f, and the unfolding rate, k_u, as in equation 3.1 and the unfolded fraction, χ_{unf}, by the equilibrium measurements as in equation 3.2, we calculated the folding rate and the unfolding rate at $T = 22°\text{C} : k_f \simeq 240,000\ \text{s}^{-1}$ and $k_u \simeq 80,000\ \text{s}^{-1}$.

$$\tau_{\text{obs}}^{-1}(T) = k_f(T) + k_u(T). \tag{3.1}$$

$$\chi_{\text{unf}}(T) = k_u(T)/(k_f(T) + k_u(T)). \tag{3.2}$$

The observed, folding and unfolding rates as the function of temperature are plotted in figure 3–6. Fitting these data into the two-state transition model (1.2), we obtained the activation energy barriers: 27±1 kJ/mol for the folding and 76±5 kJ/mol for the unfolding process.

3.3 TrpZips

3.3.1 Background

Recently the experimental studies on the formation of the secondary structure, β-sheet, become fruitful [63, 64, 65] mainly because short peptide aggregation has been avoided [66].

The TrpZips we studied are 12-residue polypeptide chains, designed by Cochran and coworkers [67]. These peptide chains are able to form β-hairpins which are the minimum units of β tertiary structures. They are shorter than what people considered to be the lower limit ($\sim$ 20–30 residues) for a stable tertiary fold and are good models for exploring the dynamics of the unit structures of protein molecules. The sequences and structural information of TrpZip1–3 are listed in table 3–1

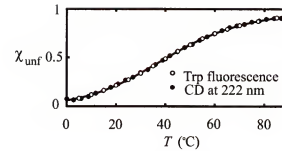


Figure 3–4: Equilibrium unfolding of TrpCage. Fraction unfolded in 15 mM phosphate, pH 7.0, as derived from 222 nm circular dichroism data and by Trp fluorescence data (266 nm excitation), with fit to two-state transition (solid curve) indicating $\Delta H \simeq 48.6$ kJ/mol and $\Delta S \simeq 155$ kJ/mol-K for unfolding. Trp-cage synthesis by solid-phase method (FMOC) with reverse-phase HPLC purification (C-18 column).

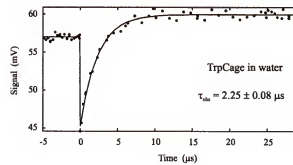


Figure 3–5: Kinetics of TrpCage. Samples are at 50 μM in 0.1 M phosphate buffer, pH 7.0. $T(t < 0) = 18.6$°C, $T(t > 0) = 25.0$°C. The solid line ($t > 0$) is the fit to a single exponential function.

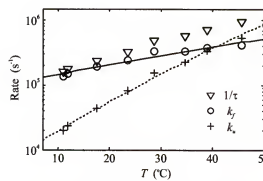


Figure 3–6: Arrhenius plot of TrpCage. The triangles are the observed relaxation rates ($1/\tau_{obs}$), Circles are folding rates (k_f) and the plus signs are unfolding rates (k_u). The solid and dotted lines are the fits to the two-state model.

Table 3–1: Sequence and structure information of TrpZips.

Materials	Amino Acid Sequence	Structure type
TrpZip1	SWTWE GNKWT WK	Type II' turn
TrpZip2	SWTWE NGKWT WK	Type I' turn
TrpZip3	SWTWE pNKWT WK	Type II' turn

The TrpZip β-hairpins are stabilized by the tryptophan-tryptophan cross strand pairs. Figure 3–7 shows the CD spectra of TrpZip1 which is characteristic for the peptide models with two identical fluorophores interacting with each other, two tryptophans for each TrpZips. These three TrpZips are different from each other on the two amino acid residues at the positions, No.6 and No.7, which are underlined in table 3–1.

The thermal unfolding transition curves for TrpZip1–3 are reversible and highly cooperative as shown in figure 3–9, figure 3–12, and figure 3–17. The thermal unfolding and sedimentation analyses are listed in table 3–2 [67]. The equilibrium constants ($K = k_f/k_u$) and unfolding fractions (χ_{unf}) as the function of temperature are plotted in figure 3–8.

Table 3–2: Thermal unfolding parameters of TrpZip peptides.

Parameter	TrpZip1	TrpZip2	TrpZip3
T_m (K)	323.0 ± 0.3	345 ± 0.1	351.8 ± 0.2
ΔH_m $(cal\, mol^{-1})$	$10,790 \pm 120$	$16,770 \pm 60$	$13,020 \pm 70$
ΔS_m $(cal\, mol^{-1}\, K^{-1})$	33.4	48.6	37.0
ΔC_p $(cal\, mol^{-1}\, K^{-1})$	231.4 ± 4	281 ± 2	195 ± 2

3.3.2 TrpZip1

TrpZip1 was prepared in 50 mM potassium phosphate buffer pH 7.0. The fluorescence and CD measurements of TrpZip1 are shown in figure 3–9. The Trp fluorescence intensity of TrpZip1 decreases as temperature increases by 2%/°C which is slightly less than that for the free Trp (−2.2%/°C). Similar to the TrpCage, this indicates that the unfolding of TrpZip peptide chains causes a weak increase of fluorescence intensity which balances the decrease the intrinsic

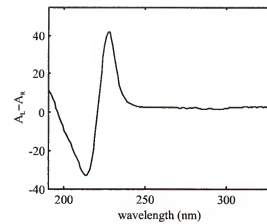


Figure 3–7: CD spectrum of TrpZip1. The TrpZip1 sample was at 50μM in 20 mM phosphate buffer pH 7.0.

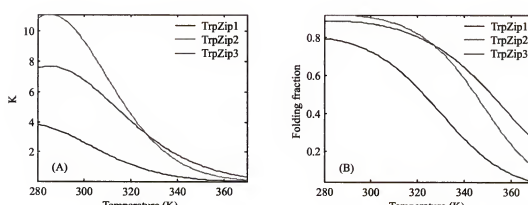


Figure 3–8: Equilibrium constants (A) and folding fraction (B) of TrpZip1–3 as the function of temperature according to the thermal unfolding and sedimentation analysis of Cochran's.

Trp fluorescence as the temperature increases. The CD measurement suggests the melting temperature (T_m) [3] of TrpZip1 is around 323K which is consistent with Cochran's measurement.

The kinetics of TrpZip1 is probed by the temperature jump spectrometer. The kinetic data at the room temperature, 22°C, is shown in figure 3–10. The relaxation time constant, τ_{obs}, is 4.6±0.18 μs. The calculated folding and unfolding time constants are $\tau_f = 6.69 \pm 0.32$ μs and $\tau_u = 19.9 \pm 0.61$ μs respectively.

The temperature dependence of the folding rates is plotted in figure 3–11. By fitting the data into the two-state model, we obtained the folding activation energy barrier, $E_a = 22.2 \pm 1.4$ kJ/mol.

3.3.3 TrpZip2

The fluorescence and CD measurements on TrpZip2 are shown in figure 3–12. The CD measurement suggests that TrpZip2 is stabler than TrpZip1 in the potassium phosphate buffer pH 7.0 and its melting temperature is around 345K.

There is no significant relaxation process observable in the temperature jump experiments at room temperature for the TrpZip2 in potassium phosphate buffer because of the very small portion of the unfolded molecules at this temperature. Adding the denaturant, Guanidine Hydrochloride, GdnHCl, helps increase the population of the unfolded state and shifts the melting temperature down as shown in figure 3–12(B). We measured the relaxation time constants τ_{obs} for the TrpZip2 solutions with varying concentrations of GdnHCl and extrapolated the time constant at [GdnHCl]=0 M: $\tau_{\text{obs}} = 1.77 \pm 0.3$ μs, as shown in the

[3] The melting temperature (T_m) of protein folding is the temperature at which the folding rate k_f and the unfolding rate k_u are equal, the populations of the folded and the unfolded molecules are equal to each other.

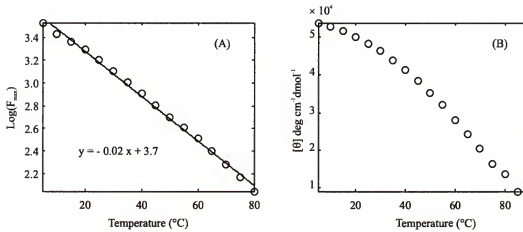


Figure 3–9: Equilibrium measurement of TrpZip1. (A) Trp fluorescence and (B) CD signal at λ_{max}=229 nm as a function of temperature. The solid line in (A) is the linear fit for $\ln(F_{max})$ vs T.

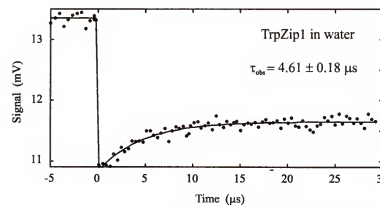


Figure 3–10: Kinetics of TrpZip1 at room temperature. The solid line (t>0) is the fit to a single exponential function.

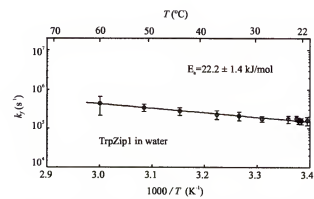


Figure 3–11: Arrhenius plot of TrpZip1. The solid line is the linear fit using the two-state model.

inset of the figure 3–13. The calculated folding and unfolding time constants are $\tau_f = 1.99 \pm 0.38$ μs and $\tau_u = 16.3 \pm 2.76$ μs respectively.

The folding rates as a function of temperature are plotted in figure 3–11. By fitting the data into the two-state model, we obtained the folding activation energy barrier $E_a = 15.1 \pm 4.9$ kJ/mol.

The relaxation rate of TrpZip2 at the room temperature is only dependent on the final temperature, not on the initial temperature or the amplitude of the temperature jump as shown in figure 3–15. However, at high temperature, ~50°C, there is a clear trend that the relaxation time constant decreases as the amplitude of the temperature jump increases as indicated in figure 3–16 (A). Figure 3–16 (B) shows how the relaxation process of TrpZip2 deviates from the single exponential function at T_{final} =50°C ($\Delta T = 20$°C).

3.3.4 TrpZip3

TrpZip3 is even stabler at room temperature and the melting temperature is about 352K. The equilibrium measurements are shown in figure 3–17. Adding 3 M denaturant GdnHCl shifts T_{m} down to ~333K as shown in figure 3–12(B).

The inset of figure 3–18 shows that the relaxation rate is slowed down by the denaturant GdnHCl, and the extrapolated relaxation time constant for TrpZip3 in water is 0.56±0.18 μs. The calculated folding and unfolding time constants respectively are $\tau_f = 0.82 \pm 0.13$ μs and $\tau_u = 5.79 \pm 0.83$ μs at 21°C.

3.4 Discussion

Two state model and ultrafast folders. The two-state model for protein folding assumes that the folded and unfolded states are separated by an activation energy barrier. The energy barrier is significant so that the probability for the barrier crossing events is much smaller than that for the molecules to transform their conformations within the unfolded state, as demonstrated in figure 3–19(A). So the relatively slow barrier crossing event becomes the rate-limiting step and

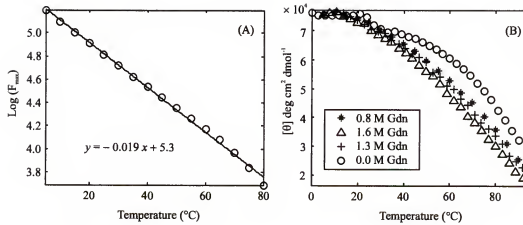


Figure 3–12: Fluorescence (A) and CD measurement at 229 nm (B) of Zip2 as the function of temperature. The solid line in (A) is the linear fit of the $\ln(F)$ versus T.

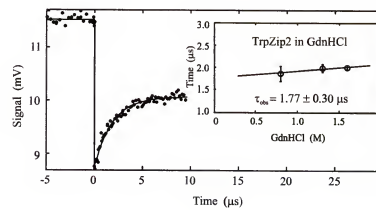


Figure 3–13: Kinetics of TrpZip2 at room temperature. The inset is the denaturant concentration [GdnHCl] dependence of the relaxation time constants.

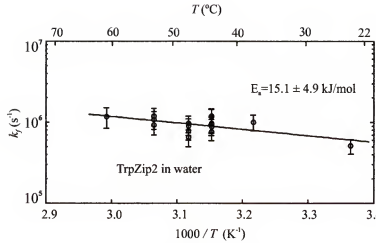


Figure 3–14: Arrhenius plot of TrpZip2. The solid line is the linear fit using the two-state model.

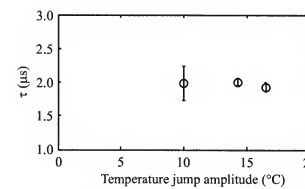


Figure 3–15: Temperature jump amplitude dependence of the relaxation rates at T_{final}=24°C. TrpZip2 sample is at 50 μM in 20 mM potassium phosphate buffer with 1.6 M GdnHCl, pH 7.0.

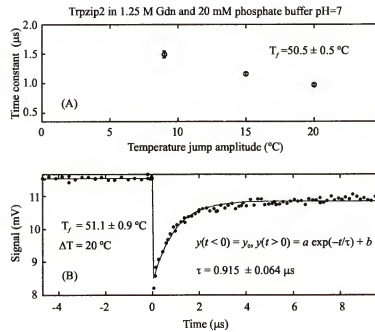


Figure 3–16: "Strange" behavior of TrpZip2 kinetics. (A) Temperature jump dependence of the relaxation rate at 50°C. (B) Relaxation kinetics at 50°C with a temperature jump amplitude of 20°C.

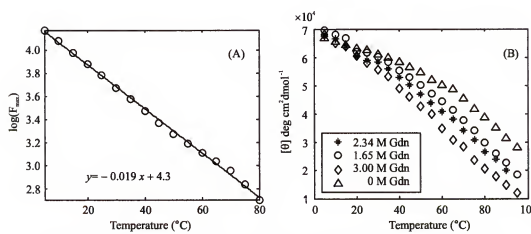


Figure 3–17: Fluorescence (A) and CD measurement (B) of TrpZip3.

the folding rate should only depend on the height of the energy barrier as in equation 1.2. In this way, one should expect a single exponential function for the kinetics although the reverse is not necessarily true for all probes [68]. The mini-protein TrpCage and the β-hairpin TrpZip1 are two-state folders although their folding activation energy barriers are only slightly higher than that associated with the viscosity of solvent which limits the rate of the loop formation ($\sim$17 kJ/mol). Their folding rates fit into the linear Arrhenius plot ($\ln k_f = \ln k_0 - \frac{E_a}{k_B}\frac{1}{T}$) quite well and are only dependent on the final temperatures.

One the other hand, when the energy barrier is negligible as shown in figure 3–20(B), the barrier crossing rate becomes comparable with the equilibration rates of molecules around the local free energy minima. There is no single transition event which dominates the folding so we observe a non-single exponential kinetics. TrpZip2 has an activation energy barrier about 15 kJ/mol and folds at $(2\ \mu s)^{-1}$ which is approaching the diffusion time limit for the loop formation [29, 30]. Its kinetic data at 50°C clearly deviate from the single exponential function as one should expect for low energy barrier folding (3–16 (B)).

In addition, figure 3–16 (A) shows that the relaxation time constant of TrpZip2 distinctly decreases as the function of the temperature jump amplitude. Similar phenomena were observed by other fast kinetic studies [28, 69]. Yang *et.al.* observed a very rapid transition event ($\sim 2\mu s$) preceding the rate-limiting step during the refolding of the five-helix bundle protein λ_{6-85}. They assigned this fast kinetics as the molecular equilibration time within the unfolded state due to the change of the free energy profile upon T-jump. As shown in figure 3–20, this model agrees with their experimental data very well. In the case of TrpZip2, the equilibration may overlap with the folding process on the time scale $\sim \mu s$. Consequently we see stretched exponential kinetics with the time constant dependent on the size of the perturbation.

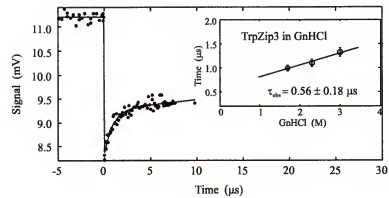


Figure 3–18: Kinetics of Zip3 with 3 M GdnHCl at room temperature. The inset is the GdnHCl concentration dependence of the relaxation rates. The solid line in the inset is the linear fit producing the relaxation time constant at 0 M GdnHCl, $0.56 \pm 0.18 \mu s$.

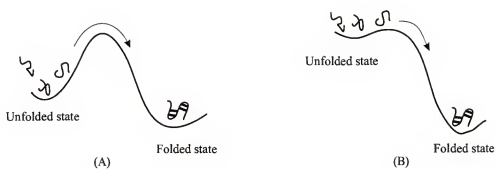


Figure 3–19: Folding with a barrier (A) and without a significant barrier (B).

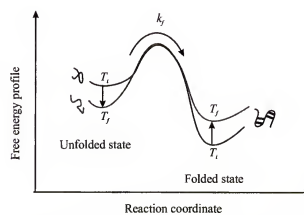


Figure 3–20: Free energy profile before and after a temperature jump.

Why do these small protein and peptides fold fast? TrpCage folds 100 times faster than what the contact order model predicts [54, 55]. The most obvious reason would be the small sizes of these molecules ($\leq$20 residues). The other possibility is that TrpCage mainly contains α-helix and 3_{10}-helix which are local structures and form in a few hundred nanoseconds [9]. Diffusion-collision model predicts that the stabilization of local secondary structures will lead to the acceleration of folding [70, 4]. Many helical proteins have been found to fold fast: Wild type λ repressor folds with folding rate 4900 s^{-1} [71, 72]; Engrailed Homedomain (En-HD), a 61-residue helical protein, folds in 27 μs; Cytochrome c folds in 500 μs; B-domain of staphylococcal protein A folds with an extrapolated folding rate 120,000 s^{-1}[7]. In contrast, the nucleation mechanism proposes that the formation of the tertiary contacts are the determinants of the folding rate [73]. The stabilization of the local structure speeds up the folding only if the native-like secondary structures are formed in the transition state ensemble. The relationship between secondary, tertiary interaction and folding rate has been investigated experimentally [74, 75]. It was shown that secondary structure interactions play a negligible role in determining the folding kinetics. Meanwhile strong and specific hydrophobic core formation plays a dominant role [76] that will be discussed more in detail in Chapter 4.

Fast folding kinetics and distributed simulation. We compare our T-jump kinetic data with folding and unfolding kinetic data from molecular dynamics simulation by V. Pande's group as shown in table 3–3. Among them, the simulation and experimental results on TrpCage and TrpZip1 are in good agreement with each other.

3.5 Conclusion

TrpCage is the smallest stable protein-like molecule and the TrpZips are among the shortest peptide chains which form β-hairpins. They all fold on the

Table 3–3: Comparison between the results of the T-jump experiments and the simulations by V. Pande's group.

Materials	Rates	Experimental Results	Simulation Estimates
TrpCage	k_f	4.10 μs	1.5–6.9 μs
TrpZip1	k_f	6.69±0.32 μs	5–7 μs
TrpZip1	k_u	19.9±0.61 μs	9–11 μs
TrpZip2	k_f	1.95±0.40 μs	3–6 μs
TrpZip2	k_u	19.3±3.57 μs	14–20 μs
TrpZip3	k_f	0.83±0.13 μs	100–300 μs
TrpZip3	k_u	5.88±0.84 μs	9–12 μs

microsecond time scale approaching the physical time limit which is set by the diffusion process of chains. TrpCage, made of a few helices and with a small hydrophobic core, provided us an ideal model both for experiment and theory. Meanwhile, the kinetic results of the TrpZip series confirmed the prediction by the theoretical models on the formation of β-hairpins and opened up an interesting field where the limitation of the structure and kinetics is being reached. In addition, some of our T-jump kinetic results on those peptides are found to be, excitingly, in agreement with molecular dynamics simulations.

CHAPTER 4
COLLAPSE OF CYTOCHROME *C* AND ITS FRAGMENTS

4.1 Background: Collapse of Polypeptide Chains

Protein molecules fold accompanied by a fast chain contraction from an extended chain into a compact globule like collapse of polymers in a poor solvent. This overall contraction is mainly due to the hydrophobic effect: apolar residues of protein molecules tend to aggregate in the presence of water. The hydrophobic effect includes not only the energetics associated with van der Waals interactions among the hydrophobic residues of proteins, but also the energetics associated with the reconstructing water which takes place upon hydration. It is one of the main driving forces for protein folding and the most important factors stabilizing native proteins. It is also believed to be an universal event for all flexible polymer chains. Analytical studies of the chain dynamics in polymers contribute considerable insight into the folding kinetics of proteins especially at the very early stage.

Theories predicted that chain collapse is on a nanosecond time scale, but experimentally observed protein collapse happens on a very wide time range, from a few nanoseconds to seconds. Some of them couple with the formation of some native secondary or tertiary structures. For examples, apomyoglobin [77, 78, 46] collapse in less than 20 μs involving at least two native helices; IgG binding domain of 62-residue α/β protein L collapses in 10 seconds concomitant with the rate-limiting step during folding [25]. In contrast, the 90-residue helical protein barstar collapses within a few milliseconds into a compact, large hydrophobic cluster without any optically active secondary or tertiary structure [26]. Gutin [79] provided two scenarios for the protein collapse: if the overall attraction due to the hydrophobic effect dominates, folding is preceded by a rapid collapse into

a structureless compact globule. If the overall attraction is just as strong as to cooperate with the native contact formation, a two-state transition from a coil to native conformation occurs. This suggests that the highly compact conformation may increase the activation energy barrier and slow down the following folding rate. In agreement, Silow [80] showed that high viscosity solvent slows down the folding by prematurely collapsing the coil and low viscosity helps folding by stabilizing the partially folded transition state. Clearly chain collapse plays a critical role in protein folding kinetics.

4.1.1 How Fast Can A Protein Fold

The classical picture of chain dynamics is based on the notion of relaxation modes proposed by P.E. Rouse in 1953. Rouse model is the Langevin dynamics of an ideal chain which is modeled as a chain of beads of coordinates $\{\vec{r_n}\}$ and mass m_n attached by springs, as in equation (4.1)

$$m_n \frac{d^2\vec{r_n}}{dt^2} + \gamma \frac{d\vec{r_n}}{dt} + \frac{\partial H}{\partial \vec{r_n}} = \vec{\xi_n}(t) \tag{4.1}$$

Left hand side are the inertial term, the friction term and the force, the right hand side is Gaussian random force. The friction coefficient γ is equal to $k_B T/D$ where D is the diffusion constant for a monomer, T is temperature and k_B is the Boltzman constant.

For a protein, the inertial term can be neglected because the mass of amino acids, m, is smaller than γ by a factor of 10^{-13}. So the Rouse equation is written as

$$\frac{d\vec{r_n}}{dt} = -\Gamma_0 \frac{\partial H}{\partial \vec{r_n}} + \vec{\eta_n} \tag{4.2}$$

where $\Gamma_0 = D/k_B T$, $\vec{\eta_n}$ is the Gaussian distributed thermal noise, and H represents the interaction among the monomers on the chain. The Rouse chain is worm-like and its end-end distances obey Gaussian statistics. Equation of motion 4.2 can be solved by using Fourier transforms. Sparing the mathematic details, we come to

the longest relaxation time of the Rouse chain, $\tau_1 \sim 10^{-11}N^2$ s, N is the number of monomers. For a protein of length $N \sim 300$, this gives relaxation time of order $\sim$1 μs.

This Zimm Model takes into account the hydrodynamic drag force among the chain residues and the interaction between the residues and solvent molecules. It applies to the polymer in dilute solution. The first order of the relaxation time is written as equation 4.3

$$\tau_1 \simeq \frac{\eta_s R_g^3}{k_B T} \tag{4.3}$$

where η_s is the solvent viscosity, R_g is the radius gyration

$$R_g^2 = \frac{1}{N}\sum_{n,m=1}^{N} < (R_n - R_m)^2 >$$

which represents the mean square length between all the pairs of the segments on the chain. This predicts the collapse time constant of tens of nanoseconds for a chain of $\sim$100 residues at room temperature.

4.1.2 What Slow Down The Protein Collapse

The chain collapse rates predicted by analytical models are 1000$\times$ faster than what are observed in experiments. This is probably because that the conformation frustration effect and energy barriers for protein folding are neglected in these models. Recent experimental and theoretical studies on short peptide chains [39, 45, 29, 30, 59, 81, 27, 82, 83] suggest that the collapse of polypeptide chain is much like a first-order transition. The collapse rates are limited by diffusion and desolvation processes which give $\sim$1 μs for the ends of 10-residue peptide chain to diffuse together and $\sim$10 μs for protein molecules of $\sim$ 100 residues to "dry" out. However this does not agree with the millisecond collapse of proteins like barstar [26], so there must be some additional factors which govern the rate of protein collapse. To clarify these issues one needs to observe the collapse on different polypeptide chains such as wild type protein molecules and unfoldable

polypeptide chains. Wild type protein molecules may or may not collapse with the formation of native structure. Unfoldable polypeptide chains can mimic the "pure" collapse of protein molecules without forming any specific structures.

4.1.3 What Could Unfoldable Peptide Chains Tell Us

To be able to see a "pure" collapse, we adopted two fragments of 104-residue helical protein cytochrome *c*, F1-65 and F1-80. These fragments cannot form any stable native structure at room temperature without the stabilizing *C*-terminal α helix. We compared the kinetics of the collapse of these unfoldable polypeptide chains with that of the wild type protein. There were basically two controversial explanations on the collapse of cytochrome *c* and its fragments. Shastry and coworkers [24, 84] observed the initial collapse of cytochrome *c* by doing continuous-flow mixing experiments. Experimental results showed that this process has single-exponential kinetics and the Arrhenius fit of the rates yields 31 kJ/mol which is larger than the energy associated with the formation of a loop controlled by diffusion ($\sim$ 17 kJ/mol). They concluded that the protein must form some native structure during the initial collapse. However it has been challenged by Sosnick and coworkers [85] who indirectly compared cytochrome *c* and its fragments by stop-flow CD measurements. They argued that because the fluorescence signals of the protein and unfoldable fragments drop to the same level after the "burst" phase, the initial collapse process just reflect a rapid solvent-dependent readjustment of a chain rather than the formation of any partially folded state.

Our kinetic results by *T*-jump spectroscopy show that both protein and its fragments collapse with the time constants $\sim$10 μs. The Arrhenius fits indicate that there are significant energy barriers limiting the collapse rate of these fragments in aqueous solvent *p*H 7.0. The barriers are apparently dependent on the chain length.

4.2 Materials

Horse heart cytochrome c is a single-domain small heme protein for electron transport. It has 104 amino acid residues, three major alpha and two minor 3_{10} helices as shown in figure 4–1. It contains a fluorophore tryptophan at the position No.59 and a fluorescence quencher, heme, at His18. By monitoring the Trp fluorescence intensity, the distance between the tryptophan and the heme can be monitored by the Förster energy transfer principle described in chapter 2. This also tells us the degree of compactness of the molecule. This protein has served as a fast folding and collapse model for numerous equilibrium and kinetic folding studies so far [86, 87, 88, 89].

Unfoldable fragments of cytochrome *c*, F1-65 and F1-80 peptides, are prepared by cyanogen bromide cleavage of ferricytochrome *c* (Sigma Chemical Co.) at the unoxidized Met65 and Met80. They were purified by reverse phase high performance liquid chromatography on a C18 column and the molecular weights are verified by mass spectrometry. Both of F1-65 and F1-80 are unable to fold into the native structure because of the lacking C-terminal α helix. The collapse of these unfoldable polypeptides represents that of a heteropolymer chain without coupling with any native structure formation.

Figure 4–1: Ribbon diagram of cytochrome *c*.

4.3 Methods and Results

4.3.1 Circular Dichroism

The circular dichroism spectra of F1-65 and F1-80 don't have the typical absorption for α-helix which has the minimum at 222 nm, like that for the wild type folded cytochrome c at the room temperature (25°C). They are more like the thermally unfolded cytochrome c molecules at 85°C or random coils. See figure 4–2 (A)

4.3.2 Fluorescence

The intensity of the tryptophan fluorescence of cytochrome c and its fragments is dependent on temperature and the distance between the tryptophan and the heme. This is related to the size of the molecules.

The protein polypeptide chains expand during the unfolding process in the presence of denaturant, GdnHCl. The Trp fluorescence will consequently increase because Trp is moving away from its quencher, the heme. Figure 4–2 (B) shows that the Trp fluorescence increases 35% and the average Trp-heme distance changes 10% upon the addition of 0.5 M GdnHCl. Figure 4–2 (C) shows the thermally expansion of the F1-65 chains. The relative fluorescence signal to the reference, Trp-containing peptide, is used to cancel out the GdnHCl concentration and temperature dependence of the tryptophan fluorescence itself.

4.3.3 Temperature Jump Spectroscopy

The samples we used for the temperature jump kinetic studies are at around 100 μM. Cytochrome c was in 1.5 M GdnHCl and 0.1 M phosphate buffer, pH 7.0. The presence of GdnHCl helps destabilize the cytochrome c molecules to some extent so that the relaxation process becomes observable in the T-jump experiments. F1-65 and F1-80 were in 0.1 M phosphate buffer and pH 7.0 without any denaturant.

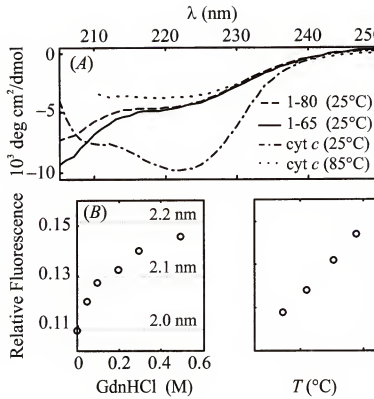


Figure 4–2: (A) Equilibrium far-UV circular dichroism of cytochrome *c* and its fragments, F1-65 and F1-80 in 0.1 M phosphate buffer, *p*H 7.0. Fragments has much less secondary structure content than that of the intact folded protein, comparable to that of the thermally denatured protein at 85°C. (B) Trp fluorescence of F1-65 fragment, relative to a reference peptide representing the Trp-containing region (residues 55-63, sequence KGITWKEET) of intact cytochrome *c*. Rising emission at higher denaturant (GdnHCl) levels indicates reduced Förster quenching of Trp-59 fluorescence by the heme, as a result of chain expansion. Dotted lines show average fluorescence of a Gaussian ensemble of random coils having the indicated mean heme-Trp distance. (C)Temperature dependence of Trp fluorescence of F1-65 in 0.1 M phosphate, *p*H 7.0, relative to the reference peptide.

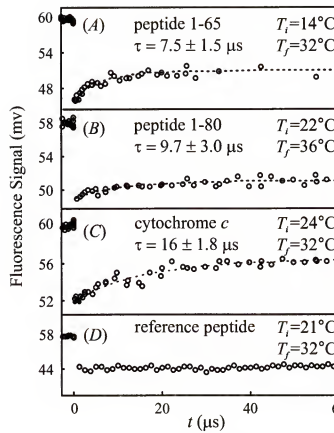


Figure 4–3: Relaxation processes observed in tryptophan fluorescence following nanosecond temperature jump at t=0, for (A) F1-65, (B) F1-80 fragments of cytochrome c in 0.1 M phosphate buffer, pH 7.0, (C) wild-type cytochrome c in 1.5 M GdnHCl, 0.1 M phosphate buffer, pH 7.0, and (D) Trptophan-containing reference peptide (KGITWKEET) in 0.1 M phosphate buffer pH 7.0

Figure 4–3 shows that temperature jump triggered the fast relaxations of F1-65 (A), F1-80 (B) and cytochrome c (C). Figure 4–3 (D) shows the T-jump response of reference peptide fluorescence. The T-jump at t=0 causes an abrupt drop of fluorescence intensity due to the intrinsic negative T-dependence of Trp fluorescence and also causes a quick shifting of the systems' equilibrium to a new one. As a result, the population of the molecules in the compact or the extended state will be redistributed as a function of time. By fitting data to the single exponential function we obtained the time constants: $\tau_{\rm obs} \simeq 25$ μs at 22°C, $\tau_{\rm obs} \simeq 8$ μs at 32°C for F1-65 and $\tau_{\rm obs} \simeq 10$ μs at 32°C, $\tau_{\rm obs} \simeq 16$ μs at 32°C for F1-80. Wild type cytochrome c is a little slower $\tau_{\rm obs} \simeq 16$ μs at 32°C. All of them seem to collapse on the time scale ~ 10 μs. There is a slight correlation between the collapse rate and the length of the chain.

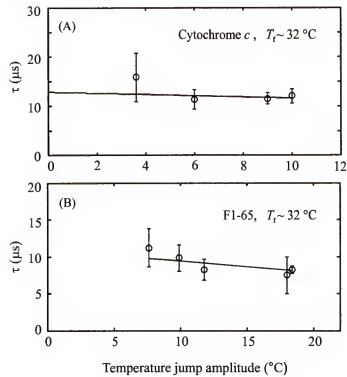


Figure 4–4: Relaxation time constants vs the temperature jump amplitude and the linear fits. (A) cytochrome c: slope=-0.12 ± 0.51 μs/°C, intercept=12.76 ± 4.52 μs. (B)F1-65: slope=-0.15 ± 0.14 μs/°C, intercept=10.95 ± 2.42 μs.

We found that relaxation times, τ_{obs}, are not sensitive to the temperature jump amplitude which is shown in figure 4–4. However, they strongly depend on the final temperature in the range 17–38°C. That suggests that the two-state transition model holds for the collapse kinetics and the collapse rate is controlled by the energy barrier crossing event. We obtained the activation energy barriers $E_a \simeq 57 \pm 7$ kJ/mol for F1-65, $E_a \simeq 68 \pm 13$ kJ/mol for F1-80 and $E_a \simeq 26 \pm 6$ kJ/mol for cytochrome c by fitting data to the Arrhenius law: $\ln \tau = \frac{\Delta E_a}{RT} + \ln \tau_0$.

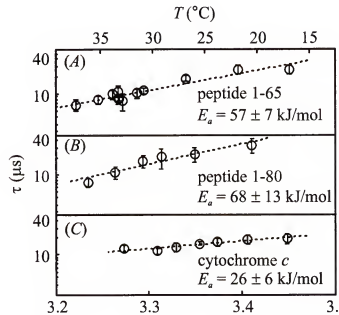


Figure 4–5: Relaxation time constant vs temperature and Arrhenius fits (dashed line) for (A)F1-65 and (B) F1-80 fragments of cytochrome c in pH 7.0 phosphate buffer, and (C) wild-type cytochrome c in 1.5 M denaturant (GdnHCl), pH 7.0.

In addition, the presence of the low concentration (less than 2 M) of denaturant (GnHCl) does not effect the collapse rate of cytochrome c very much. As shown in figure 4–6, we can compare the the collapse rate of cytochrome c in 1.5 M GdnHCl directly to those for fragments in the phosphate buffer.

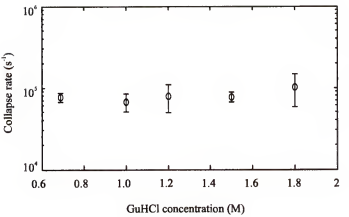


Figure 4–6: GdnHCl concentration dependence of the collapse rate of wild type cytochrome *c*.

4.4 Discussion

The mechanism of the collapse has not been fully understood. Collapse itself could be very fast $\sim \mu$s or it may happen concomitantly with the rate-limiting step. How different protein collapse is from homopolymers, how the collapse effects the whole folding process and what is the relationship between the collapse and the steric structure of proteins is crucial in solving the protein folding problem.

4.4.1 Collapse of Polypeptides and Folding Time Limit

Different from the homopolymers, the components of protein, amino acids can be cataloged into at least two groups : hydrophobic(HP) and hydrophilic(HL) according to the degree of aversion to aqueous solvents. HP residues tend to segregate into a interior core while HL residues stay on the external surface. This kind of conformations are not accessible to homopolymers although the ensemble of compact conformations of a homopolymer is gigantic and also increases as k^n with chain length n ($k \simeq 1.74$) [90]. The Zimm model predicts the relaxation time constants for the non-interacting chains of the size as F1-65 and F1-80: $\tau_1 \simeq 21$ ns and $\tau_1 \simeq 24$ ns respectively in water at 22°C according to equation 4.3. Nevertheless the observed time constants for these unfoldable peptide chains are about a thousand times slower ($\sim 10\mu$s). The hydrophobicity-induced chain

contraction involved conformation entropy and enthalpy loss may play a role in it. Since the collapse of the wild type protein cytochrome c is on the same time scale as its unfoldable analogs, we may conclude that the microsecond peptide chain collapse is not controlled by forming native structure. A recent study on the refolding of metastable compact cytochrome c [91] shows that at the low viscosity limit the folding rate approaches a certain value $\sim$ (10 μs)$^{-1}$ instead of infinity. It suggests that the folding kinetics is temporally limited by the factors such as the intra-chain interactions among the elements of protein molecules in addition to the diffusion process. Furthermore, one should expect that the energy barrier associated with the intrachain interaction increases as the length of chain increases. The T-jump studies here provide us a piece of evidence that the energy barrier for the collapse of F1-80 is higher than that for the F1-65 by $\sim$19% with the chain length $\sim$23% longer.

On the other hand, Bieri and coworkers [82] directly measured the intramolecular chain diffusion of polypeptides by the determination of the triplet-triplet energy transfer rates and found out that the loop formation of the shortest peptides (N=3) has a time constant of 20 ns which seems to be consistent with what was predicted for homopolymer chains.

Protein chain collapse involves the diffusion, desolvation and intramolecular interactions which are dependent on individual proteins. This explains the diversity of collapse mechanisms and time scale flexibility.

4.4.2 Collapse and Folding Efficiency

There are a lot of debates on the roles chain collapse plays in protein folding, like whether or not the collapse induced compactness enhances the formation of secondary structures, α-helix and β-sheet [31, 32, 92, 93]. Maritan *et. al.* [33, 34] demonstrated that helix is the optimal shape of compact strings due to packing efficiency and also forming helices is in favor of fast folding. This not only contributes

insight into the relationship of the packing of the whole molecule and the secondary structure formation but also provides the evidence that nature adopts some special criteria like space occupying and folding temporal efficiencies in the selecting amino acid sequences for protein molecules.

4.5 Summary

We fully resolved the collapse of unfoldable polypeptides and wild type cytochrome c by the laser induced temperature jump spectrometer. By comparing the kinetic data on the foldable and unfoldable peptide chains we found all of them collapse in ~ 10 μs. These fragments experience significant energy barriers. The barriers are larger than expected for diffusion due to nonspecific contacts. Interactions inside protein molecules limit the collapse rates by roughening the free energy landscape and increasing the activation energy barrier.

Our study demonstrates that the kinetics is not as fundamental as sequence and steric structure in protein folding because it can be varied to some extent by changing of external conditions like temperature, pressure or concentration of denaturant [94, 95, 96]. Combining the accurate kinetic measurements by T-jump with the specific structure probes like circular dichroism can fruitfully help us understand fast protein folding dynamics.

CHAPTER 5
CONCLUSION

We constructed a nanosecond laser induced temperature jump spectrometer. It is of great utility in identifying local structure and tertiary contact formation kinetics of protein molecules on time scale of nanosecond to millisecond.

The kinetic studies on TrpCage show that this 20-residue helical protein molecule folds in a highly cooperative way with a folding rate $k_f \simeq 240,000\ s^{-1}$, thousands times faster than what the contact order model predicted ($k_f \simeq 100\ s^{-1}$). There is no evidence showing the existence of any intermediate state lying between its folded and unfolded states. Arrhenius fit to k_f versus T data indicates that the folding encounters a very weak activation energy barrier $E_a = 27 \pm 1$ kJ/mol which slightly exceeds that associated with the viscosity of the solvent $E_a \sim 17$ kJ/mol which sets the time limit ($\sim 0.2\ \mu$s) for ends of 20-residue chain to diffuse together.

We also investigated the TrpZips series 12-residue peptide chains. They form β-hairpins which are the minimum units of β tertiary structure. The folding rates of β-hairpins varies in the range of 1–10 μs. The folding and unfolding time constants are consistent with what molecular dynamic simulations predicted.

These small protein and short peptide chains received much attention not just because they are two-state folders and fold fast but also because this study is where experiment and theory meet each other with exciting agreement in their results.

Cytochrome c of 104-residue is also a two-state folder but it folds accompanied by fast chain collapse preceding the rate-limiting step. There have been lots of debates on the fast chain collapse mainly on the issues such as whether or not protein molecules forms native-like structure upon collapse. This distinguishes

them from the synthetic polymers. Another questin is what sets the collapse rates. We address these questions by studying the foldable wild type cytochrome c and its unfoldable fragments F1-65 and F1-80. We found that both kinds of peptide chains collapse at time constants around 10 μs, and experience activation energy barriers between 20–70 kJ/mol. These activation energy barriers are probably due to the interactions between the residues on the peptide chain instead of forming any specific native-like structures.

Understanding how fast protein folds is one of the center issues of the protein folding problem. Fast probes like the temperature jump spectrometer makes this study possible. More and more fast kinetic data on proteins will help people reveal the whole story on protein folding.

APPENDIX A
AMINO ACIDS

Table A–1: List of 20 amino acids

Amino acids	Three-letter code	One-letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	lle	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

APPENDIX B
LABVIEW PROGRAMS

B.1 Program for alignment

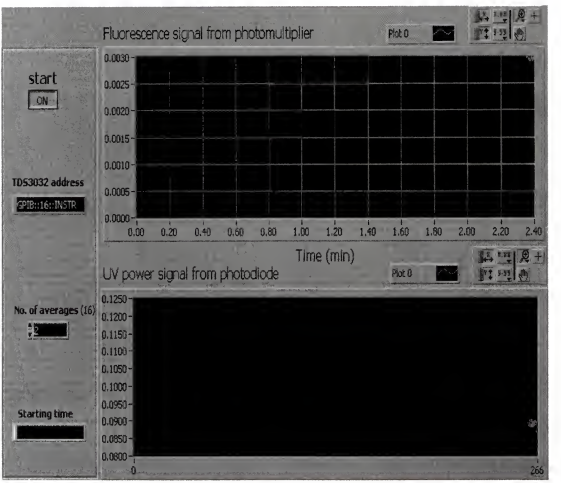


Figure B–1: Front panel of *alignmentsavedata*

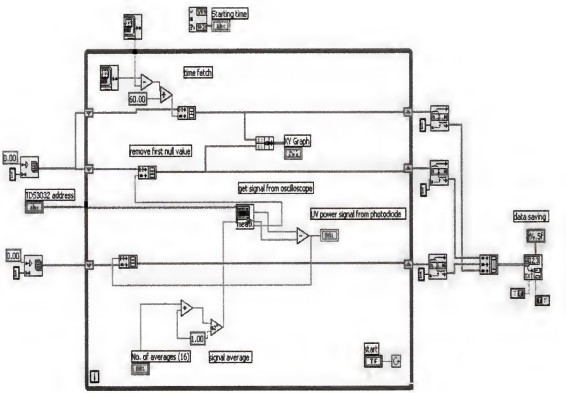


Figure B–2: Block diagram of *alignmentsavedata*

B.2 Program for data collection

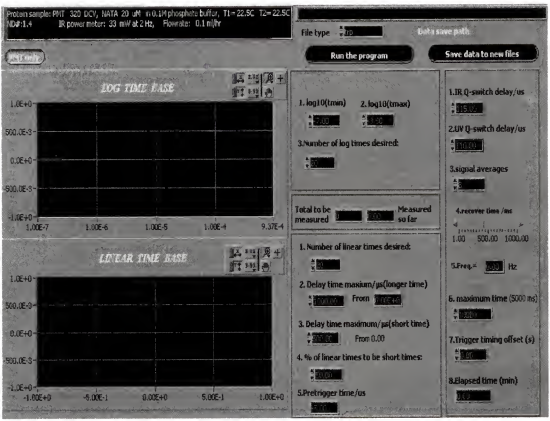


Figure B–3: Front panel of *tjumpRMS.vi*

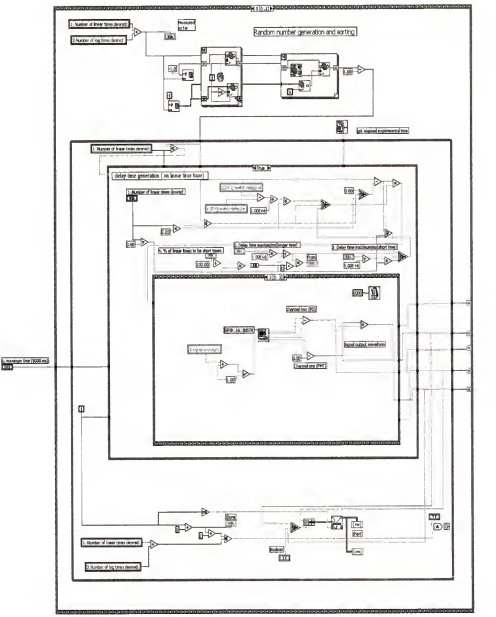
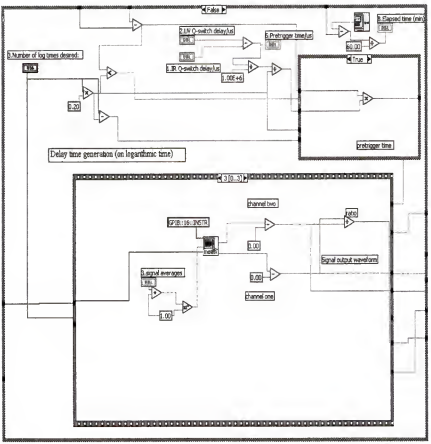


Figure B–4: Block diagram of *tjumpRMS.vi*–1



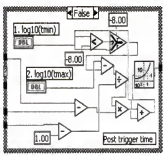


Figure B–5: Block diagram of *tjumpRMS.vi*–2

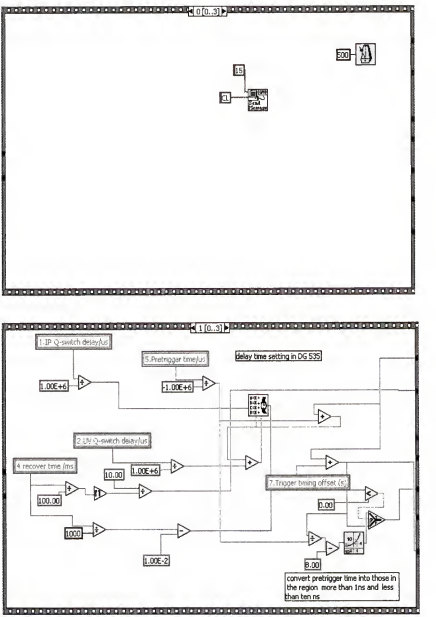


Figure B–6: Block diagram of *tjumpRMS.vi*–3

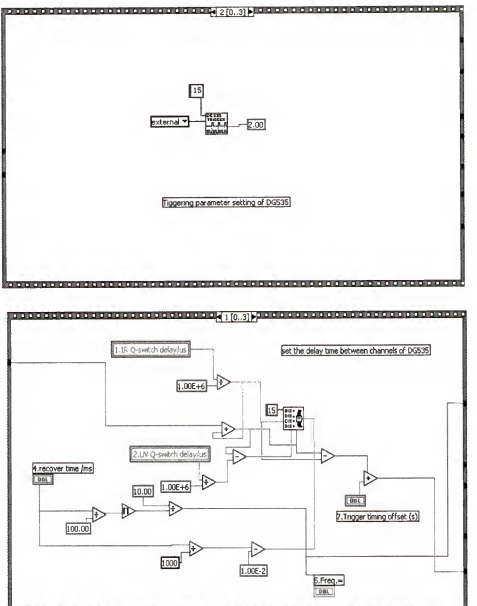
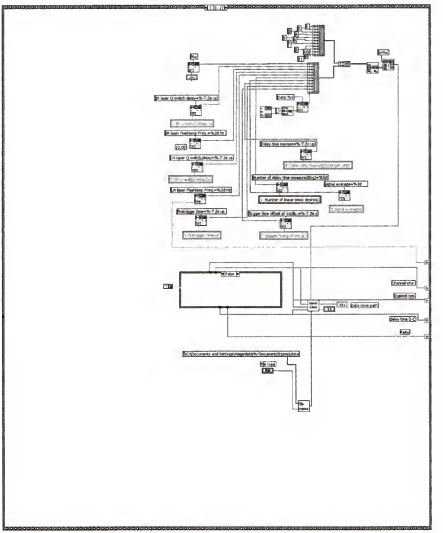


Figure B–7: Block diagram of *tjumpRMS.vi*–4



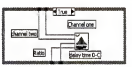


Figure B–8: Block diagram of *tjumpRMS.vi*–5

APPENDIX C
MAPLE PROGRAM FOR THE TRPZIP1 KINETIC DATA ANALYSIS

```
> restart; with(plots):


> y:=[6.3,8.16,5.51,5.70,6.38,6.48]:  N:=nops(y):

> x:=[22.8,19.6,23.1,24.5,21.4,22.3]:  T0:=-22:

> s_y:=[1.2,2.55,0.84,1.22,1.52,1.19]:

> s_x:=[.3,.52,1.61,1.66,1.2,.98]:

> comp:=proc(a,b)
> description "compare two 2d points according to the first
> coordinate";

> if a[1]<b[1] then return true;

> else return false; end if

> end proc:
> xy:=sort([seq([x[i]+T0, y[i]],i=1..N)],comp); #sort data points
into

> the up order of x

> x:=[seq(xy[i][1],i=1..N)]:  y:=[seq(xy[i][2],i=1..N)]:
```

$xy := [[-2.4, 8.16], [-.6, 6.38], [.3, 6.48], [.8, 6.3], [1.1, 5.51], [2.5, 5.70]]$

```
> chi2:=0:

> for i from 1 to N do

> chi2:=chi2+(y[i]-a-b*x[i])^2/(s_y[i]^2+b^2*s_x[i]^2);

> end do:      #chi2:=simplify(chi2);

> chi:=minimize(chi2,a=5..7,b=-10..10,location);
```

```
> con:=chi[2][1][1]; chi0:=chi[1];
```

$$\chi := .2862324527, \{[\{b = -.5390217386,\ a = 6.691620021\},\ .2862324527]\}$$

$$con := \{b = -.5390217386,\ a = 6.691620021\}$$

$$\chi 0 := .2862324527$$

```
> eqn:=2*chi0=chi2:

> a2:=solve(subs(con[1],eqn)):  b2:=solve(subs(con[2],eqn)):
> a0:=(a2[1]+a2[2])/2; da:=abs(a2[1]-a2[2])/2; b0:=(b2[1]+b2[2])/2;
> db:=abs(b2[1]-b2[2])/2;
```

$$a0 := 6.691620020$$

$$da := .3134760970$$

$$b0 := -.5461458158$$

$$db := .1891066274$$

```
> T:=22+T0:  tau:=subs(con,a+b*T); dtau:=sqrt(da^2+T^2*db^2);
```

$$\tau := 6.691620021$$

$$dtau := .3134760970$$

```
> a3:=[6.38,7.01]; b3:=[-0.35,-0.73]; # max and min values on
2*chi0
> contour curve
```

$$a3 := [6.38,\ 7.01]$$

$$b3 := [-.35,\ -.73]$$

```
> da_1:=abs(a3[1]-a3[2])/2; db_1:=abs(b3[1]-b3[2])/2;
```

$$da\_1 := .3150000000$$

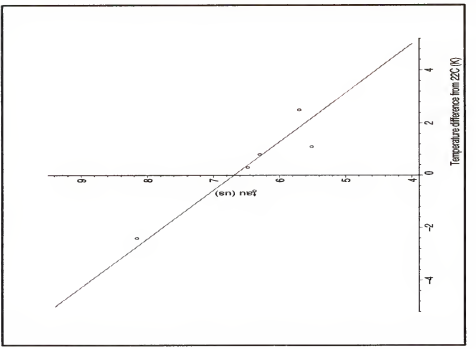
$$db\_1 := .1900000000$$

```
> tau_1:=subs(con,a+b*T); dtau:=sqrt(da_1^2+T^2*db_1^2);
```

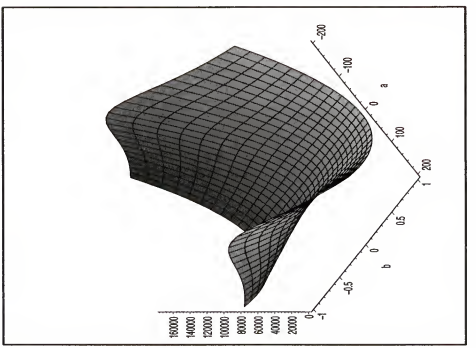
$$tau\_1 := 6.691620021$$

$$dtau := .3150000000$$

```
> plot([xy,[seq([x,subs(con,a+b*x)], x=-5..5)]],
> symbol=CIRCLE,style=[point,line],labels=["Temperature difference
from
> 22C (K)","tau (us)"], labeldirections=[horizontal,vertical]);
```




```
> plot3d(chi2, a=-200..200,b=-1..1, axes=framed);
```



```
> contourplot(chi2,a=6..7.5,b=-1.2..-0.1,contours=[1.01*chi0,1.5*chi0,2
> *chi0,1+chi0]);
```

REFERENCES

[1] C.B. Anfinsen. Principles that govern folding of protein chains. *Science*, 181:223–230, 1973.

[2] P.L. Privalov and G.I. Makhatadze. Contribution of hydration to protein folding thermodynamics. 2. the entropy and gibbs energy of hydration to protein folding. *J. Mol. Biol.*, 232:660–679, 1993.

[3] S.E. Jackson. How do small single-domain proteins fold? *Folding & Design*, 3:R81–R91, 1998.

[4] R.L. Baldwin and G.D. Rose. Is protein folding hierarchic? i. local structure and peptide folding. *Trends Biochem. Sci.*, 24:26–33, 1999.

[5] R.L. Baldwin and G.D. Rose. Is protein folding hierarchic? ii. folding intermediates and transition states. *Trends Biochem. Sci.*, 24:77–83, 1999.

[6] J.K. Myers and T.G. Oas. Mechanisms of fast protein folding. *Annu. Rev. Biochem*, 71:783–815, 2002.

[7] J.K. Myers and T.G. Oas. Preorganized secondary structure as an important determinant of fast protein folding. *Nature Structure Biology*, 8:552–558, 2001.

[8] R.H. Callender, R.B. Dyer, R. Gilmanshin, and W.H. Woodruff. Fast events in protein folding: The time evolution of primary processes. *Annu. Rev. Phys. Chem.*, 49:173–202, 1998.

[9] W.A. Eaton, V. Muñoz, P.A. Thompson, E.R. Henry, and J. Hofrichter. Kinetics and dynamics of loops α-helices, β-hairpins and fast-folding proteins. *Acc. Chem. Res.*, 31:745–753, 1998.

[10] R.B. Dyer, F. Gai, and W.H. Woodruff. Infrared studies of fast events in protein folding. *Account of Chemical Research*, 31:709–716, 1998.

[11] V. Muñoz, E.R. Henry, J. Hofrichter, and W.A. Eaton. A statistical mechanical model for β-hairpin kinetics. *Proc. Natl. Acad. Sci. USA*, 95:5872–5879, 1998.

[12] V. Muñoz, P.A. Thompson, J. Hofrichter, and W.A. Eaton. Folding dynamics and mechanism of β-hairpin formation. *Nature*, 390:196–199, 1997.

[13] G.S. Jas, W.A. Eaton, and J. Hofrichter. Effect of viscosity on the kinetics of α-helix and β-hairpin formation. *J. Phys. Chem. B*, 105:261–272, 2001.

[14] I.K. Lednev, A.S. Karnoup, M.C. Sparrow, and S.A. Asher. α-helix peptide folding and unfolding activation barriers: A nanosecond uv resonance raman study. *J. Am. Chem. Soc.*, 121:8074–8086, 1999.

[15] C. Huang, J.W. Klemke, Z. Getahun, W.F. Degrado, and F. Gai. Temperature-dependent helix-coil transition of an alanine based peptide. *J. Am. Chem. Soc.*, 123:9235–9238, 2001.

[16] C. Huang, Z. Getahun, Y. Zhu, J.W. Klemke, W.F. Degrado, and F. Gai. Helix formation via conformation diffusion search. *Proc. Natl. Acad. Sci. USA*, 99:2788–2793, 2002.

[17] A. Ansari, S.V. Kuznetsov, and Y. Shen. Configurational diffusion down a folding funnel describes the dynamics of dna hairpins. *Proc. Natl. Acad. Sci. USA*, 98:7771–7776, 2001.

[18] G. Bonnet, O. Krichevsky, and A. Libchaber. Kinetics of conformational fluctuations in dna hairpin-loops. *Proc. Natl. Acad. Sci. USA*, 95:8602–8606, 1998.

[19] B. Nölting. *Protein folding kinetics, Biophysical methods.* Springer, New York, 1999.

[20] Y. Duan and P.A. Kollman. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science*, 282:740–744, 1998.

[21] C.D. Snow, B. Zagrovic, and V.S. Pande. The trp cage: Folding kinetics and unfolded state. *Nature Structure Biology*, 9:425–430, 2002.

[22] U. Mayor, C.M. Johnson, V. Daggett, and A.R. Fersht. Protein folding and unfolding in microseconds to nanoseconds by experiment and simulation. *Proc. Natl. Acad. Sci. USA*, 97:13518–13522, 2000.

[23] V.S. Pande, I. Baker, J. Chapman, S.P. Elmer, S. Khaliq, S.M. Larson, Y.M. Rhee, M.R. Shirts, C.D. Snow, E.J. Sorin, and B. Zagrovic. Atomistic protein folding simulations on the submillisecond time scale using world distributed computing. *Biopolymer*, 68:91–109, 2003.

[24] M.C.R. Shastry and H. Roder. Evidence for barrier-limited protein folding kinetics on the microsecond time scale. *Nature Structure Biology*, 5:385–392, 1998.

[25] K.W. Plaxco, I.S. Millett, D.J. Segel, S.Doniach, and D. Baker. Chain collapse can occur concomitantly with the rate-limiting step in protein folding. *Nature Structure Biology*, 6:554–556, 1999.

[26] V.R. Agashe, M.C.R. Shastry, and J.B. Udgaonkar. Initial hydrophobic collapse in the folding of barstar. *Nature*, 377:754–757, 1995.

[27] M.S. Cheung, A.E. Garcia, and J.N. Onuchic. Protein folding mediated by solvation: Water expulsion and formation of the hydrophobic core occur after the structure collapse. *Proc. Natl. Acad. Sci. USA*, 99:685–690, 2002.

[28] W.Y. Yang and M. Gruebele. Folding at the speed limit. *Nature*, 423:193–197, 2003.

[29] S.J. Hagen, J. Hofrichter, A. Szabo, and W.A. Eaton. Diffusion-limit contact formation in unfolded cytochrome c: Estimating the maximum rate of protein folding. *Proc. Natl. Acad. Sci.*, 93:11615–11617, 1996.

[30] S.J. Hagen, J. Hofrichter, and W.A. Eaton. Rate of intrachain diffusion of unfolded cytochrome *c*. *J. Phys. Chem. B*, 101:2352–2365, 1997.

[31] D.P. Yee, H.S. Chan, T.F. Havel, and K.A. Dill. Does compactness induce secondary structure in proteins? a study of poly-alanine chains computed by distance geometry. *J. Mol. Biol.*, 241:557–573, 1994.

[32] H.S. Chan and K.A. Dill. Polymer principles in protein structure and stability. *Annu. Rev. Biophys. Biophys. Chem.*, 20:447–490, 1991.

[33] A. Maritan, C. Micheletti, and J.R. Banavar. Role of secondary motifs in fast folding polymers: A dynamical variational principle. *Phys. Rev. Lett.*, 84:3009–3012, 2000.

[34] A. Maritan, C. Micheletti, A. Trovato, and J.R. Banavar. Optimal shapes of compact strings. *Nature*, 406:287–290, 2000.

[35] P.G. de Gennes. *Scaling concepts in polymer physics*. Cornell University Press, Ithaca and London, 1979.

[36] M. Doi and S.F. Edwards. *The theory of polymer dynamics*, volume 73 of *International series of monographs on physics*. Oxford Science Publication, 1986.

[37] D. J. Brockwell, D. A. Smith, and S. E. Radford. Protein folding mechanisms: new method and emerging ideas. *Current Opinion in Structure Biology*, 10:16–25, 2000.

[38] N.A.J.V. Nuland, V. Force, J. Balbach, and C.M. Dobson. Real-time nmr studies of protein folding. *Acc. Chem. Res.*, 31:773–780, 1998.

[39] C.M. Jones, E.R. Henry, Y. Hu, C.K. Chan, S.D. Luck, A. Bhuyan, H. Roder, J. Hofrichter, and W.A. Eaton. Fast events in protein folding initiated by nanosecond laser photolysis. *Proc. Natl. Acad. Sci. USA*, 90:11860–11864, 1993.

[40] J. Hofrichter, E.R. Henry, A. Szabo, L.P. Murray, A. Ansari, C.M. Jones, M. Coletta, G. Falconi, M. Brunori, and W.A. William. Dynamics of the

quaternary conformational change in trout hemoglobin. *Biochemistry*, 30:6583–6598, 1991.

[41] C.R. Cantor and P.R. Schimmel. *Biophysical chemistry: Techniques for the study of biological structure and function*, volume 2. W.H. Freeman and Company, San Francisco, 1980.

[42] M. Eigen. Fast elementary steps in chemical reaction mechanism. *Pure and Appl. Chem..*, 6:97–115, 1963.

[43] M. Eigen and K. Kustin. The study of very rapid reactions in solution by relaxation spectrometry. *ICSU Review of World Science*, 5:97–115, 1963.

[44] M. Gruebele, J. Sabelko, R. Ballew, and J. Ervin. Laser temperature jump induced protein refolding. *Acc. Chem. Res.*, 31:699–707, 1998.

[45] S.J. Hagen and W.A. Eaton. Two-state expansion and collapse of a polypeptide. *J. Mol. Biol.*, 301:1019–1027, 2000.

[46] R. Gilmanshin, S. Williams, R.H. Callender, W.H. Woodruff, and R.B. Dyer. Fast events in protein folding: Relaxation dynamics of secondary and tertiary structure in native apomyoglobin. *Proc. Natl. Acad. Sci. USA*, 94:3709–3713, 1997.

[47] Bengt Nölting. Temperature-jump induced fast refolding of cold-unfolded protein. *Biochemical and Biophysical Communications*, 227:903–908, 1996.

[48] K.P. Murphy. *Protein structure, stability, and folding*, volume 168 of *Methods in molecular biology*. Humana Press, Totowa, NJ, 2001.

[49] W.O. Wray, T. Aida, and R.B. Dyer. Photoacoustic cavitation and heat transfer effects in the laser-induced temperature jump in water. *Appl. Phys. B*, 74:57–66, 2002.

[50] S.E. Jackson. Folding of chymotrpsin inhibitor 2: 1. evidence for a two-sate transition. *Biochemistry*, 30:10428–10435, 1991.

[51] P.L. Privalov. Thermodynamic problems of protein-structure. *Annu. Rev. Biophys. Biophys. Chem.*, 18:47–69, 1989.

[52] K. Dill. Theory for the folding and stability of globular proteins. *Biochemistry*, 24:1501–1509, 1985.

[53] E.I. Shakhnovich and A.V. Finkelstein. Theory of cooperative transitions in protein molecules 1. why denaturation of globular protein is a first-order phase transition. *Biopolymer*, 28:1667–1680, 1989.

[54] D.E. Makarov, C.A. Keller, K.W. Plaxco, and H. Metiu. How the folding rate constant of simple, single-domain proteins depends on the number of native contacts. *Proc. Natl. Acad. Sci. USA*, 99:3535–3539, 2002.

[55] D.E. Makarov and K.W. Plaxco. The topomer search model: A simple, quantitative theory of two-state protein folding kinetics. *Protein Science*, 12:17–26, 2003.

[56] K.W. Plaxco, S. Larson, I. Ruczinski, D.S. Riddle, E.C. Thayer, B. Buchwitz, A.R. Davidson, and D. Baker. Evolutionary conservation in protein folding kinetics. *J. Mol. Biol.*, 298:303–312, 2000.

[57] N.V. Dokholyan, L. Li, F. Ding, and E.I. Shakhnovich. Topological determinants of protein folding. *Proc. Natl. Acad. Sci. USA*, 99:8637–8641, 2002.

[58] S.B. Ozkan, K.A. Dill, and I. Bahar. Fast-folding protein kinetics, hidden intermediates and the sequential stabilization. *Protein Science*, 11:1958–1970, 2002.

[59] L. Qiu, C. Zachariah, and S.J. Hagen. Fast chain contraction during folding: "foldability" and collapse dynamics. *Phys. Rev. Lett.*, 90:168103 1–4, 2003.

[60] L.L. Qiu, S.A. Pabit, A.E. Roitberg, and S.J. Hagen. Smaller and faster: The 20-residue trpcage protein folds in 4 μs. *J. Am. Chem. Soc.*, 124:12952–12953, 2002.

[61] J.W. Neidigh, R.M. Fesinmey, and N. H. Anderson. Designing a 20-residue protein. *Nature Structure Biology*, 9:425–430, 2002.

[62] C. Simmerling, B. Strockbin, and A.E. Roitberg. All-atom structure prediction and folding simulations of a stable protein. *J. Am. Chem. Soc*, 124:11258–11259, 2002.

[63] E. de Alba, M.A. Jiménez, M. Rico, and J.L. Nieto. Conformational investigation of designed short linear peptides able to fold into β-hairpin structures in aqueous solution. *Fold. Des.*, 1:133–144, 1996.

[64] T. Kortemme, M. Ramirez-Alvarado, and L. Serrano. Design of a 20-amino acid, three-stranded beta-sheet protein. *Science*, 281:253–256, 1998.

[65] M. Ramirez-Alvarado, F.J. Blanco, and L. Serrano. *De novo* design and structural analysis of a model beta-hairpin peptide system. *Natl. Struct. Biol.*, 3:604–612, 1996.

[66] H.J. Dyson and P.E. Wright. Defining solution conformations of small linear peptides. *Annu. Rev. Biophys. Biophys. Chem.*, 20:519–538, 1991.

[67] A.G. Cochran, N.J. Skelton, and M.A. Starovasnik. Trptophan zippers: Stable, monomeric β-hairpins. *J. Am. Chem. Soc.*, 124:14548–14549, 2002.

[68] S.J. Hagen. Exponential decay kinetics in "downhill" protein folding. *Protein: Structure, Function, and Genetics*, 50:1–4, 2003.

[69] J. Sabelko, J. Ervin, and M. Gruebele. Observation of strange kinetics in protein folding. *Proc. Natl. Acad. Sci. USA*, 96:6031–6036, 1999.

[70] M. Karplus and D. Weaver and. Protein-folding dynamics. *Nature*, 160:404–408, 1976.

[71] G.S. Huang and T.G. Oas. Structure and stability of monomeric λ repressor: Nmr evidence for two-state folding. *Biochemistry*, 34:3884–3892, 1995.

[72] R.E. Burton, G.S. Huang, M.A. Daugherty, P.W. Fullbright, and T.G. Oas. Microsecond protein folding through a compact transition state. *J. Mol. Biol.*, 263:311–322, 1996.

[73] E.I. Shakhnovich. Theoretical studies of protein-folding thermodynamics and kinetics. *Curr. Opin. Struct. Biol.*, 7:29–40, 1997.

[74] T.R. Sosnick, S. Jackson, S.W. Englander, and W. Degrado. The role of helix formation in the folding of a fully α-helical coiled-coil. *Proteins*, 24:427–432, 1997.

[75] J.K. Myers and T.G. Oas. Reinterpretation of gcn4-p1 folding kinetics: Partial helix formation precedes dimerization in coiled coil folding. *J. Mol. Biol.*, 289:205–209, 1999.

[76] E. Durr, I. Jelesarov, and H.R. Bossard. Extremely fast folding of a very stable leucine zipper with a strengthened hydrophobic core and lacking electrostatic interactions between helices. *Biochemistry*, 38:870–880, 1999.

[77] R.M. Ballew, J.Sabelko, and M. Gruebele. Direct observation of fast protein folding: The initial collapse of apomyoglobin. *Proc. Natl. Acad. Sci. USA*, 6:5759–5764, 1996.

[78] R. Gilmanshin, S. Williams, R.H. Callender, W.H. Woodruff, R.B. Dyer, and M.C.R. Shastry. Fast events in protein folding: Relaxation dynamics of secondary and tertiary structure in native apomyoglobin. *Proc. Natl. Acad. Sci. USA*, 94:3709–3713, 1997.

[79] A.M. Gutin, V.I. Abkevich, and E.I. Shakhnovich. Is burst hydrophobic collapse necessary for protein folding. *Biochemistry*, 34:3066–3076, 1995.

[80] M. Silow and M. Oliveberg. High concentrations of viscogens decrease the protein folding rate constant by prematurely collapsing the coil. *J. Mol. Biol.*, 326:263–271, 2003.

[81] P.R.T. Wolde and D. Chandler. Drying-induced hydrophobic polymer collapse. *Proc. Natl. Acad. Sci. USA*, 99:6539–6543, 2002.

[82] O. Bieri, J. Wirz, B. Hellrung, M. Schutkowski, M. Drewello, and T. Kiefhaber. The speed limit for protein folding measured by triplet-triplet energy transfer. *Proc. Natl. Acad. Sci. USA*, 96:9597–9601, 1999.

[83] M. Jacob, T. Schindler, J. Balbach, and F.X. Schmid. Diffusion control in an elementary protein folding reaction. *Proc. Natl. Acad. Sci USA*, 94:5622–5627, 1997.

[84] M.C.R. Shastry, J.M. Sauder, and H. Roder. Kinetic and structural analysis of submillisecond folding events in cytochrome *c*. *Acc. Chem. Res*, 31:717–725, 1998.

[85] T.R. Sosnick, M.D. Shlerman, L. Mayne, and S.W. Englander. Ultrafast signals in protein foldings and the polypeptide contracted state. *Proc. Natl. Acad. Sci. USA*, 94:8545–8550, 1997.

[86] A.S. Morar, A. Olteanu, G.B. Young, and G.J. Pielak. Solvent-induced collapse of α-synuclein and acid-denatured cytochrome *c*. *Protein Science*, 10:2195–2199, 2001.

[87] S. Yeh and D.L. Rousseau. Hierarchical folding of cytochrome *c*. *Nature Structure Biology*, 7:443–445, 2000.

[88] S. Yeh, S. Han, and D.L. Rousseau. Cytochrome *c* folding and unfolding: A biphasic mechanism. *Acc. Chem. Res.*, 31:727–736, 1998.

[89] S.W. Englander, T.R. Sosnick, L.C. Mayne, M. Shtilerman, P.X. Qi, and Y. Bai. Fast and slow folding in cytochrome *c*. *Acc. Chem. Res*, 31:737–744, 1998.

[90] E.I. Shakhnovich and A.M. Gutin. Implications of thermodynamic of protein folding for evolution of primary sequences. *Nature*, 346:773–775, 1990.

[91] S.A. Pabit and S.J. Hagen. Friction-limited folding of a compact denatured protein. to be submitted, 2003.

[92] N.G. Hunt, L.M. Gregoret, and F.E. Cohen. The origins of protein secondary structure effects of packing density and hydrogen bonding studied by a fast conformation search. *J. Mol. Biol.*, 241:241–225, 1994.

[93] N.D. Socci, W.S. Bialek, and J.N. Onuchic. Properties and origins of protein secondary structure. *Phys. Rev. E*, 49:3440–3443, 1994.

[94] A.P. Capaldi, C. Kleanthous, and S.E. Radford. Im7 folding mechanism: misfolding on a path to the native state. *Nature Structure Biology*, 9:209–216, 2002.

[95] H. Nguyen, M. Jäger, A. Moretto, M. Gruebele, and J.W. Kelly. Tuning the free-energy landscape of a ww domain by temperature, mutation and truncation. *Proc. Natl. Acad. Sci. USA*, 100:3948–3953, 2003.

[96] J. Karanicolas and C.L. Brooks III. The structural basis for biphasic kinetics in the folding of the ww domain from a formin binding protein: Lessons for protein design. *Proc. Natl. Acad. Sci. USA*, 100:3954–3959, 2003.

BIOGRAPHICAL SKETCH

Linlin Qiu was born in Xianyang, P.R. China, on July 24, 1971. In September 1989, she entered Shannxi University majoring in electrical and electronic communications and graduated with a bachelor's degree. In September 1993, she started the graduate school of Shanghai Jiaotong University majoring in optics and photonics and graduated with a Master of Science degree. She entered the graduate school of the University of Florida majoring in Physics, subfield biophysics in August 1999 and graduated with a Doctor of Philosophy degree in August, 2003.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Stephen J. Hagen, Chair
Assistant Professor of Physics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Haiping Cheng
Associate Professor of Physics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

David Reitze
Professor of Physics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Andrew Rinzler
Associate Professor of Physics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Christopher Stanton
Professor of Physics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Adrian E. Roitberg
Associate Scientist of Chemistry

This dissertation was submitted to the Graduate Faculty of the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 2003

Dean, Graduate School

LASER INDUCED TEMPERATURE JUMP INVESTIGATIONS OF FAST PROTEIN DYNAMICS

Linlin Qiu
(352) 846-3110
Department of Physics
Chair: Stephen J. Hagen
Degree: Doctor of Philosophy
Graduation Date: August 2003

Understanding how protein molecule fold spontaneously from a random chain into a three-dimension structure is very important in life science because it helps people to predict protein structures from their amino acid sequences, design new drugs and fix the misfolded that causes the diseases like Madcow and Alzheimer's.

To study the fast folding kinetics of proteins and polypeptide chains, we constructed a laser induced temperature jump spectrometer which is capable of identifying the local structure and tertiary contact formation of protein molecules on the time scale between 10^{-8} and 10^{-3} s with the time resolution of nanoseconds.

With the T-jump spectrometer, we carried out the ultrafast kinetics studies on a small protein TrpCage, β-hairpin series TrpZips and fast chain collapse studies on wild type cytochrome c and its unfoldable fragments F1-65 and F1-80. Our results turned out to be in good agreement with theoretical and computer simulation studies and attracted much attention in this field.